

生成AIをめぐるデータガバナンス —オープン化とエンパワーメントの緊張関係

渡辺智暁¹（国際大学）

要 旨

本稿は生成AIの学習データのガバナンスをとりあげ、日本の現在の著作権法上担保されているようかなり自由な利用の可能性が望ましいかどうかを検討する。併せて、オープン化の概念を援用しつつ学習データをかなり自由に利用できることがほかのオープン化的文脈と同様に肯定的に評価できるとは限らない理由について検討する。

生成AIの開発にあたって重要なリソースである学習データは現在の日本の著作権制度の下ではかなり自由に利用できる。オープン化の概念を手掛かりに考えるなら、これは生成AIの開発、利用にとってよいアプローチと言える。開発の費用が低いことで競争が生まれやすくなり、提供の価格が低くなることで生成AIのような汎用技術の用途探索を加速させる効果が期待できるため、社会的便益の早期実現に資する可能性や、そのような有用性をもたらす技術への進化も期待できる。

ただし、そのような考え方には、市場構造、技術の特性、リスク、倫理課題、法解釈など多くの領域の主要論点に関する前提条件や意見の上に成り立つもののようにも思われる。これら前提条件や意見が変化すると結論にも影響が及ぶ可能性があるため、学際的な議論が重要となる。

経済的な恩恵を考えた場合には開発と普及を促進することにメリットがあると思われるが、学習データとして自らの著作物が利用される人の中にはそれによって経済的な被害を受ける人もいるかも知れない点などが難点である。これを開発者からの対価の支払いや補償金制度で解決することはおそらく困難で、市場競争への影響などを考えるなら学習データの購入費が参入障壁となることも望ましくないと思われる。他に、文化の大きな変容が起こる可能性についてどのように考えるか、軍事利用などのリスクをどう考えるか、といった論点もある。

オープン化は影響力の小さかった個人などをエンパワーし、その活躍をよりよい社会などにつなげる仕組みとして機能して来た面があるが、生成AIについてはそのような構図とは大きく異なる構図も存在している。学習データとして利用される著作物を提供するのは個人や小規模集団なども含めた多数のクリエイターであり、巨大なコンテンツ産業ではない。それを利用する生成AI開発企業には巨大企業が含まれる。オープン化の推進をしてきた国際的な組織が、公正さや互酬性といった、利用のしやすさとは少し異なる規範を重視するようになってきたことは、このような変化と呼応していると思われる。

キーワード：オープン化、AI、データガバナンス、著作権

¹ 国際大学グローバル・コミュニケーション・センター（GLOCOM）教授、主幹研究員、研究部長

1. はじめに

本稿では特に近年の生成AIに関わるデータ利活用に着目しながら、データをめぐるガバナンスについて検討したい。「データ」として注目したいのは生成AIの開発にあたって「学習データ」として用いられる様々な情報資源で、イラスト、写真などの視覚芸術的な作品、報道記事などの文字による作品、プログラミングコードなど著作物を多く含む。中には権利が存在しなくなったパブリックドメインのものや、数値データなどで著作物ではない情報資源も含まれる。本稿では筆者の知識の制約があるため、主に文章と画像の生成AIを意識して議論を進める。

非常に大雑把に言えば、日本の著作権法は学習データについては、開発者はかなり自由に（個別に著作権者から許諾を得ることなく）データを学習に利用することができる法制度となっている。データの利活用によってさまざまな便益を得る可能性を考えた場合、このように自由な学習利用ができるることは、都合のよい、優れた制度であるように思われる。だが、そのように断じてよいかについては多くの疑問もある。そこで、このような制度はいわばデータ活用社会のためのデータガバナンスの一つの形として優れているのではないか、というラフな前提から出発しつつ、このような考え方の妥当性についてより細かな論点を検討するという形で本稿を進めていく。

この本稿の主題は問い合わせの形で表現すると比較的単純なものだが、データ活用社会のためのデータガバナンスの最適解と思われるものを出すことは容易ではないように思われる。検討にあたっては難所も多く、事実関係やその中長期的変動の有無についてのいくつもの前提をおいた議論になる。適切な判断のためには学際的な視野からの詳細な議論や多くの実証的な調査を必要としているように思われる。前提条件が変動すれば最適解も変更を迫られる部分も多い。そこで、様々な前提条件とその帰結について概観しつつ、「難所」としてはオープンデータなどデータ利活用を容易にすること一般と比べた場合の難しさに特に注目した議論を行うこととする。後者についてはこれを情報社会の大きな変化と結び付けて位置づける。具体的には、幅広い人をエンパワーするようなある種の「民主化」的な性質を持つことが多かった「オープン化」の取り組みが、生成AIについての議論ではむしろ特定少数の受益者の権益増大になるのではないか、という「格差拡大」「搾取」などの懸念と結びつきやすい点についてとりあげる。これらを通じて、幅広い分野の専門家がこの問題についての議論についてより精度の高い議論をするための踏み台を提供することを狙いたい。

以下ではまず検討にあたって前提とするオープン化の概念について述べ、AIの開発や普及を促進する上でも優れたアプローチであるという可能性を述べる（第2節）続いて学習データとしての情報資源利用について数多く存在する主要な論点の概観と検討を行い、様々な前提条件や判断の違いが結論を左右することになることを示し（第3節）、それを受けた政策・ガバナンスの制約や主要オプションの評価を行う（第4節）。

なお、オープン化の概念（あるいは、制約の少ない、自由なデータ利活用）は本稿で後に述べるような定義に基づくなら狭義には著作権法と密接に関係したものだが、広義には様々な法的、その他の要素とも関わりがある。本稿では筆者の知見の制約と、紙幅を考慮して著作権を中心とした考察を行う。法的要素に着目するなら、個人情報保護法やプライバシー権などはデータガバナンスを考慮する上では少なくとも同等に重要だが、それらをとりあげた考察については別の機会に譲ることにしたい。

2. 情報資源利用のコストとオープン化

2. 1. オープン化と取引コスト、利用コスト

オープン化を推進する立場に立ってシンプルに考えるならば、学習データの利用を極めて容易にし、生成物の利用も容易にしている現在の著作権法制度は、オープン化を推進する度合いが高いものとして評価できるだろう。だが、そもそもオープン化とは何だろうか？まずはその点について整理する。

効果の側面に着目してデータガバナンスのような文脈でのオープン化を捉えるなら、その主要な特徴は著作物の取引コストを低減させることにある²。データを利用するにあたって必要になる費用（調査・交渉などの時間も含む）を減らしてデータを使いやすくする、と言い換てもよい。取引コストが高い状態とは、例えば、他人の著作物を利用するが原理的には可能であっても、そのために許諾をもらわなければならず、許諾料の交渉をしなければならず、それに関連して利用の様態・目的などについての条件について合意を形成しなければならず、それらの前段として著作権者を探し出さなければならない、というようなコストがかかる状態だ。著作物の利用やそのほかの情報資源の利用の文脈では、このような、利用者側のコストが特に問題になりやすい³。

取引コストを含め、著作物を利用するためには必要なコストが高ければ、人は他人の著作物を利用することを断念しやすい。取引コストの概念には通常含まない商品の価格（著作物で言えば利用料・許諾料）も、高額であれば当然利用を断念する人が増える。このようなコストが極めて低いものには「フリー素材」「著作権フリー素材」などと一般に言われているものがあり、取引コストだけでなく、実際の利用料も無料のものが多くあるが、社会でこうした情報資源が広く使われていることひとつをとっても、取引コストが著作物利用に与えている影響を推しあかることができるだろう。オープンデータと呼ばれるような政府保有データを一般開放する取り組みにおいても、利用条件を簡素なものにし、無料で、利用にあたっての申請・承認手続きなども不要にする、ということが広く行われる。データの利用価格をゼロまたは限界費用まで下げるに、中小企業を中心に利用者が 1,000%-10,000%増加する

² 著作権制度と取引コストの関係については、例えば田中（2018）また、近年の著作権政策の中で取引コスト低減を意識したものとしては、例えば未管理著作物裁判制度がある（文化庁著作権課、2025）。立法に至るまでの主な議論の経緯については小山（2023）。当該論文でも言及されている知財戦略本部下の構想委員会コンテンツ小委員会デジタル時代における著作権制度・関連政策の在り方検討タスクフォースはその中間とりまとめ内で、取引コストの削減を重視し、「コンテンツ産業・市場の発展にとどまらない、デジタル・エコノミー発展上の重要課題ということになる。」（p.14）とした上で政策対応に向けた検討を促している。（知財戦略本部構想委員会コンテンツ小委員会デジタル時代における著作権制度・関連政策の在り方検討タスクフォース、2021）。

³ 取引コスト全般の議論の中では、契約を締結した後にそれを執行するコストも取引コストの種類として挙げられることが多い。これはデータの利用者側よりも提供者側が負担する傾向が強いコストと言えるだろう。一般に、取引コストは著作物や情報資源の「利用者」側だけが負担するものとは限らず、「提供者」側の負担になる部分もある。例えば、自分の著作物を一般に公開すると意に沿わない使われ方をすることが予期されるが、それらをやめさせるのも、事前審査制にして個別に検討するのも大変なので、そうであれば一般には公開しない、といった状況を考えることができるが、これは著作物の提供者側の執行に関わる取引コストが高い程生じやすい状況であると言える。

と報告した調査もある⁴。

また、生成AIの文脈では学習データの一種として用いられることがあるソフトウェアのコードについても、同様に取引コスト低減事例がある。いわゆるオープンソースソフトウェアだ。このようなソフトウェアは公衆一般に対して利用許諾があらかじめ付与されており、通常無償で提供されている。これにより取引コストを含めた利用に関わるコストを低く抑えることができる。必ずしも意識されているわけではないが、OSやウェブブラウザを含めオープンソースソフトウェアは様々な領域で広く活用されているところだ。ほかにも先に少し触れた「オープンデータ」のように政府などが保有する情報資源を（著作物に限らず）広く一般に開放し、あらかじめ利用許諾を与えておくような取り組みも世界中でなされ、様々な利用例が生まれている⁵。取引コストの観点からは、著作権法は、情報資源を原則として著作作者に許諾を得なければ利用できないものとする法律であるが故に、取引コストの原因となっている。これを「パブリック・ライセンス」によって克服することが、オープン化の一つの典型的な手法になっている。パブリック・ライセンスというのは、公衆一般に対して許諾を与えるタイプのライセンスである。通常のライセンスが相対契約でライセンスを与える側と受ける側がそれぞれ特定の人であるようなものと対照的に、誰でも個別の交渉や連絡なしに利用できる点で取引コストを抑制する効果がある。

ただし、取引コストを低減させる手段はそのようなライセンスにとどまるものではない。データやソフトウェアが検索しやすいこと、それがどのような用途に適合しているかがわかりやすいこと、ライセンスもその内容がわかりやすいこと、ほかのパブリック・ライセンスとの互換性があること、など単に許諾を与えることを超えた様々な性質や取り組みが取引コストを低減させる効果をしばしば持つ。（ライセンスだけでは取引コスト低減ができない例として、すべてのデータがパブリック・ライセンスによってかなり自由に利用できるようになっているが、そのライセンスの種類が多種多様であるような状況を想定することができる。このような状況下で複数のデータを組み合わせるような形での利用が可能かどうかは、利用したい個々のデータのライセンスの条件の間に矛盾などがないかどうか、複数のライセンスの条件を同時に満たせるかどうかなどによる。単一のライセンスのみを考慮すればよい場合と比べて格段に検討のコストが増すことになる。単にパブリック・ライセンスが付与されているだけではなく、互いに組み合わせやすいライセンスの下に提供されているデータを探す、というより難しい探索を強いられることもありうるだろう。）

また、法律の細かな規定によっても取引コストは大きく左右されることがある。例えば何が著作物に該当し、何が著作物に該当しないかについて、わかりやすく多くの人が判断できるような基準を備えた著作権法は、本稿の文脈では取引コストが低いと言える。法律を細かく理解しなければ判断できない、専門家に相談しなければ判断できない、あるいは裁判をしてみないとわからない、というような規定ぶりであるとすると、取引コストは高い。この取引コストの高さがもたらす帰結は、例えば、ある古い文書の保存状態がよくないためにデジタル化を図りたいが、果たしてその行為（複製行為、デジタル化したファイルを公開する場合には公衆送信可能化の行為）が合法であるかを検討するにあたって、著作権法の改正と共に

⁴ de Vries, et al. (2011)

⁵ オープンデータの利用例については、例えば、デジタル庁 (2025).

に異なる性質の著作物の保護期間がどのように変化してきたかを確認し、果たして古文書の著作権の保護期間が満了しているのか、そうでないのかを判断するための時間がかかる、というような形でデジタルアーカイブの構築のコストを増大させることもある⁶。

あるいは、ウィキペディアにある人物の略歴やある会社の沿革が投稿された場合などにも影響を与える。そのような投稿が、その人物を扱った別のウェブサイトの内容とよく似ていたり、会社の公式ウェブサイトに記載されている沿革とよく似ていたりした場合に、偶然にしては似すぎているからそれら外部のサイトに依拠して作成され、投稿されたものだろう、と考えられるとして、ではどの程度類似していたらその投稿は著作権侵害になるのか、どの程度であれば問題ないのか、といったことは自明ではない。自明ではないので、ウィキペディア日本語版であれば気が付いた人たちの間で議論が行われ、念のため削除する、と言ったことが起こり得る。人によってこの意見が分かれる場合には議論がまとまりにくくなったり、安全のために、幅広く（それほど類似性が高くないと考えられるものであっても）削除されたりすることもありうる。このようなプロセスを一步引いて眺めれば、著作物の利用と、著作物ではないデータの利用の境界線が明確な法の規定ぶりが仮にあったとすれば節約できていたであろうコストが、こうした情報資源の利用に際して発生しているということである。

学習データの利用により関係が深いところでは、準拠法の問題があるだろう。どのような国の法廷で訴えられるとどのような国の法律が適用され、その結果としてどのような行為が著作権法やそのほかの法律上問題となるのかは、明快とは程遠い状況にあるように思われる⁷。

情報資源の利用しやすさを左右するのは、必ずしも取引コストと利用料だけではない点も述べておくべきだろう。例えばあるデータは取得してから利用するまでに加工を必要とするため、取得までにかかる取引コストが低くても利用までにかかるコストが高い、ということがありうる。より具体的には、データのフォーマットやスキーマなどはこの利用コストを左右する要因になりやすいと言えるだろう。そのような加工コストが高い例として、地理空間系のデータを入手した際に、位置を示す情報について住所や緯度経度などのよく用いられている情報の代わりに独自のラベリングが施されており、その定義も独自のものであり、よくある地番情報や緯度経度情報には簡単に置き換えられず、従ってほかのデータと組み合わせた分析ができない、といった場合を考えてみるとわかりやすいだろう。

日本での認知度はやや限定的ではあるが、学術研究などが生み出す研究データの共有に関する FAIR 原則と呼ばれるものがある⁸。この原則は、2014 年にオランダで開催された

⁶ 数藤& 橋本 (2019). からそのコストを伺い知ることができる。

⁷ 文化審議会法制度小委員会 (2024) もこの点については明快なガイドラインを提供できていないようと思われる。デジタルヒューマニティーズ分野ではこの不明瞭さが研究の遅れにつながっているという指摘もある(Samberg, et al, 2023)。Trimble(2025)は著作権法の制限規定・例外規定と呼ばれるものが国際的に統一されていないことなどにより、国境をまたぐ形で行われる行為の何が適法で何が適法でないかが不明瞭である点を強調している。福岡(2022)は、ウェブサイトのデータの収集に関する訴訟の文脈で、外国の裁判所で適用される準拠法について「予測可能性が低い」と形容している。

⁸ NBDC 研究員 (大波純一、八塚茂、信定知江、箕輪真理、三橋信孝、畠中秀樹) (2019), Force11 (2023), Hahnel, et al. (2024), 池内 (2025).

ワークショップで、より大量のデータを活用したデジタル時代の科学研究を推進する方法のひとつとして考案されたもので、FAIR の 4 文字に対応する 4 つの大原則をおいている。それらは発見しやすさ、アクセスのしやすさ、相互運用性、再利用のしやすさ、の 4 点で、3 点目にデータの相互運用性を挙げている点は、利用コストがこの点によって大きく左右されることのひとつの反映と見ることができよう。生成 AI の文脈でも、このような利用コストは相当に高い場合があるものと思われる。例えば ABEJA 社が日本語コーパスのデータセットを作成するにあたって、元となる Common Crawl のデータをどのように処理したかについて公開しているが、そこからはウェブ上で公開されている文字データの中から、AI の学習に適したものを見出し、学習に使えるフォーマットに整える手順が高度な専門性や高価な設備へのアクセスを必要とするものであることが伺える⁹。なお、FAIR 原則についての解説資料には、あるデータセットがそもそもどのようなものであるかについての情報（メタデータ）が十分に細かいことの重要性など、取引コストについても参考になる点が盛り込まれている。

2. 2. オープン化とライセンス、利用条件（狭義のオープン化）

以上のような効果面に着目した場合のオープン化は、ある情報資源の取引コストや利用コストが減って、利用がしやすくなること、と言い換えることができるだろう。これはオープン性を量的なもの、程度を指し示すものとして捉え、低いレベルのオープン性から高いレベルのオープン性へ変化することを「オープン化」と呼ぶような考え方だと言える。だが、情報資源のオープン性については、もうひとつ非常に強い影響力を持っている考え方がある。ある情報資源のオープン化は、特定の基準値を超えるレベルのオープン性を達成することを指す、というような考え方だ。これまでにとりあげたオープンソースソフトウェアやオープンデータなどの文脈ではこの意味で「オープン」という言葉が使われている。それは「グローバルに認識されている一定のレベルを満たす形で、利用に対する制約が取り除かれている」というようなことを意味していることが多い。このような意味でのオープン性の度合いに近づくことは、それが結果として十分なレベルに達しなければ、「オープン化」とは呼べない、ということになる。これはフリーとか Libre と自由を意味する語で形容されることもある。FLOSS (Free Libre Open Source Software) など同義語を併記する形で形容されることもある。

ここでグローバルに認識されている「一定のレベル」は、複数の文書で定義がされており、分野などによって異なる定義が参照されているが、おおよその意味は共通している。その内容を見ると、商業利用や改変を含めて利用が可能である、個別の許諾を取得する手続きなしに利用が可能である、利用の条件として要求されるのは著作者へのクレジットや、作成された改変物を他人に提供する場合に元の著作物と同じように自由に利用できるものとして提供すること、という 2 点程度にとどまっている、というようなものになっている¹⁰。また、

⁹ Common Crawl から作る大規模日本語コーパスとその前処理 (Mixtral 8x7B を語彙拡張継続事前学習 Part2) , ABEJA Tech Blog

¹⁰ 主要な定義として、Open Knowledge Foundation による Open Definition, Open Source Initiative による Open Source Definition, 特定の組織によるものではないが、ウィキペディア系プロジェクトで参照される Definition of Free Culture Works, フリーソフ

これは著作権を中心とした考え方で、データベース権、特許権などに及ぶことがありうるが、そのほかの法的な権利や許諾は必ずしも含んでいない¹¹。含まれない権利・許諾の例として、パブリシティ権と呼ばれることがあるような権利が挙げられる。ある肖像写真がこのような意味での「オープン」である時、撮影した人の著作権に関する許諾は与えられており、撮影者に個別に許諾をもらうことなく改変や商業利用ができるだろう。だが、被写体となっている人は、自分の似姿などをを利用して顧客誘引力を発生させることについて、許諾を与えていないかも知れない。これは実質的に肖像写真の商業利用を大きく制限することになるが、肖像写真はそのような制限があっても「オープン」と形容されることがある。

このように「オープン性の度合いの増加としてのオープン化」と「一定の基準を満たすこととしてのオープン化」2つの用法があることでややこしさがあるものの、前者は広義の、後者は狭義のオープン化だと考えることができる。

本稿の主題からはやや外れることになるので踏み込んでは扱わないが、モデルのウェイトがオープンになっている LLM や、そのほかの情報がオープンになっているオープンソースの AI が存在する¹²。OpenAI 社の社名も、元来はそのような意味でのオープン性を意識したものであった¹³。オープンの形容はしばしば利用者に訴求力を持つことなどから濫用され、批判が起こることにもなるが AI の文脈でもそれは同じでこうした性質を持つ LLM をオープンソース、オープンなどと称し批判されることがある¹⁴。

2. 3. 決定プロセスのオープン化、オープン化全般の特徴

情報資源の利用しやすさ以外にも「オープン」と形容されることが多い事象に、参加性や透明性の高い意思決定プロセスがある。たとえばオープン性を重んじるウィキペディアのようなプロジェクトでは、百科事典の編纂に誰でもが参加できるようになっている。個々の

トウェア財団による Free Software Definition が挙げられる。ライセンスの性質についての定義と、ライセンスの下に提供される作品についての定義が混在しており、また、ソフトウェアを対象にしたもの、知識やデータを対象にしたもの、著作物を主に想定したもの、と対象物についても違いがあるが、主旨は共通する部分が多い。他に、より狭く学術文献などのオープンアクセスについての案として Budapest Open Access Initiative によるブダペスト宣言がある。また、Open Source Initiative により近年策定されたものとして The Open Source AI Definition がある。

¹¹ たとえば生成 AI でも時折採用されている Apache 2.0 ライセンスは、著作権（第 2 条）と特許権（第 3 条）に関する許諾を明示的に含み、商標権についての許諾は含まないことも明記されている（第 6 条）。<https://licenses.opensource.jp/Apache-2.0/Apache-2.0.html> このライセンスは ChatGPT シリーズの中でオープンウェイト版と呼ばれる GPT-OSS の 2 つのモデルに採用されている。（OpenAI, 2025a）

¹² Meta 社の Llama シリーズ、Google 社の Gemma、X 社の Grok 2.5、Open AI 社の GPT シリーズの一部、Mistral 社の同名のモデル、DeepSeek 社の同名のモデル、Alibaba 社の Qwen などがある。画像生成系では Qwen のほかに Black Forest 社の FLUX の一部、Stability AI 社の Stable Diffusion などが挙げられる。最も性能が高いモデルではないバージョンをオープンウェイトとしてリリースするケースが多い傾向にあるのは文章系、画像系とも共通だが、Llama や Stable Diffusion など例外もある。

¹³ Brockman, et al. (2015).

¹⁴ 例えば Widder, et al. (2024).

記事の執筆・編集だけではなく、執筆編集の方針策定も参加者が決めており、そればかりかいわゆる「サイト運営」にあたるような活動（トラブルを引き起こす参加者の処遇の決定や、その決定プロセスや判断基準などの方針策定など）も参加者が決めている¹⁵。非営利財団を設立し理事会を設置した際には、その理事の多くが参加者による投票で、参加者の中から選ばれていた。そこまでの参加者主導ではないサービスは、2000年代中盤にWeb2.0、UGC（User Generated Contents）などの語と共に注目された。現在でも非常に広く利用されているものを挙げるならYouTubeやX（旧称Twitter）がその代表格であろう。こうした投稿型サイトのコンテンツは、テレビ局などのコンテンツが局の責任と判断によって編成されることと対照的に、利用者によって作成・投稿されるものであり、運営主体による個別・事前の承認も存在しないままにコンテンツが公開されている。

このようなプロセス面でのオープン性は、情報資源利用のオープン性と全く別のものではなく、共通性がある。すなわち、特定少数の者の管理・コントロール下にある度合いが低いものをオープンであると形容し、従来よりも多数、または不特定の者に対して影響力を行使する余地を作り出すことは、オープン化と呼ばれる。オープンデータやオープンソースソフトウェアのような情報資源であれば、その複製や改変などを含めた利用の決定について、著作権者はかなりの部分利用者側に判断を委ねている。参加性の高い意思決定プロセスは、参加性の低いプロセスよりも、参加者に対して意思決定に対する影響力を付与している度合いが高い。他にオープンイノベーションという語も似た含みを持つ。この語は社外の主体と共同で何かの価値を創出する取り組みに幅広く適用される傾向があるが、社内のリソース以外のリソースを活用し、また、社内のメンバー以外のメンバーに影響力を付与することになる点、本稿で述べている広義のオープン化の性質を持っていることが多いと言えるだろう。

オープンイノベーションの概念が幅広い事象に適用されることからも想像できるように、オープン化は情報資源の活用や情報資源の開発のためのコラボレーションや意思決定の領域に限って存在しているわけではない。例えばプラットフォームのオープンAPIには似た効果がある。一時期のTwitterがそうであったように、サービスを通じてユーザーがやり取りするデータをAPI経由で取得可能にすることで、そのデータを別の形で処理・表示するようなユーザー向けのソフトウェアをTwitter社以外の主体が多く作成できるようにすることは、オープン化の一形態と言える。

オープン化は政策を通じて強制されることもある。たとえば情報通信政策の分野には古くから、ネットワークインフラの「アンバンドリング」と呼ばれる政策があるが、これはある事業者が保有する通信基盤を、競合するほかの事業者に対して提供することを義務付けるもので、強制的なオープン化の政策と言える。オープン化の結果、ネットワークインフラ資源の配分や利用について、より多くの主体が影響力を行使することができるようになる点は、オープン化一般の参加性拡大と共通している。データポータビリティに関する政策もこれと類似したところがある。特定のプラットフォームへのユーザーの囲い込みを成立にくくし、プラットフォーム間でユーザーの往来が起こりやすくなる。ユーザーはプラットフォームの使い分けや切り替えがしやすくなる結果、後発の優れたプラットフォームが先

¹⁵ ウィキペディアのガバナンスの参加性の高さについては、例えば渡辺（2010）を参照。

発のプラットフォームからユーザーを奪う、といった戦略が成立しやすくなる場合がある。うまく機能すれば、プラットフォームサービス間のより激しい市場競争を通じて、サービス提供者はよりユーザーの求めるサービスを提供するインセンティブが高まり、ユーザーは自分の好みに合ったサービスを選びやすくなり、資源配分の効率化が達成できることになるだろう。携帯電話領域のナンバーポータビリティなどもよく似た性質を持っている政策だと言える。競争促進策としてオープン化を強制する政策には、あるプラットフォームが自社と他社の間で差別的な取り扱いをすることを禁じるような政策も含まれる。例えば検索結果で自社サービスと他社サービスを差別しないことがそれにあたる。

さらに抽象性の高い議論となるが、市場経済を計画経済と比較した場合に、前者はより多くの主体が資源配分の決定に影響力を行使している点でオープン性が高い経済のあり方だということができる。民主制は独裁制に比べ同じようにオープン性が高い政治のあり方だということができる。市場経済は決定が「分散」的に行われる点で情報資源のオープンな利用と似ており、民主制は意見の「集約」過程により広い人々が「参加」する点で意思決定過程のオープンなあり方と似ているとみる事もできる。

2. 4. オープンのメリットとデメリット

このようなオープン化にはどのようなメリットがあるだろうか？これについては集合知についての研究や近年の多様性についての議論に比較的わかりやすい説を見る事ができる¹⁶。すなわち、特定少数の決定によるよりも、より不特定・多数の参加や分散的決定による方が、多数の主体の情報や発想、貢献などを取り込むことができるため、より質の高い決定ができたり、より多様なニーズに応えるような資源配分が可能になったりする。例えば百科事典に掲載すべき知識を持っている人は、伝統的な百科事典の執筆者であるような学者に限られない。実務家であったり、趣味についての深い知識を持っている人であったり、主題になっている事柄に当事者として関わった人たちであったり、ほかにも様々な人が知識を持っている。ある事柄について、どの資料のどの個所に有益な記載があるかを知っている人は、やはり専門の学者のみではないだろう。また、百科事典の記事の記載内容に抜け漏れや偏りがあるかどうかは、様々な人の視点から検討した方が、特定少数の専門家が検討するよりも多くのアイディアが出るだろう。こうした分散した知識や能力をうまく集約するには、執筆・編集を広くオープンにする方がよい。オープンにすることで質の低い投稿や悪意のある虚偽やいたずらなど様々な執筆・編集も行われることになるが、これを発見・修正する役目も、広く参加者に与えることで解決できるところが多い。（というのも、例えればいたずらをしたい人は数としては少なく、それを防止・修正したい人は数としては多く、単純ないいたずらは多くの人にとって見つけやすいものだからだ。）

なお、こうしたメリットは、社会にとってのメリットとして幅広い人々が享受することもあるが、一部のメリットはオープン化を実施した主体が受けることになる。例えば百科事典編纂のオープン化は、従来型のアプローチでは決して到達できないような規模の百科事典の作成を可能にする、より効率的な編纂を可能にする、といった形でオープン化した主体に

¹⁶ 集合知の有効性に関する研究として例えば、ペイジ（2009）がある。体系性は弱いが幅広く関連した事象を集めたものにスロウイッキー（2006）がある。

メリットをもたらすことがある。執筆・編集者の確保についても参加者がやってくれるためコストを節約できることがあるだろう。自らの著作物を自由に使えるようにするクリエイターは、それによってファンが増えたり、自分が知らなかつた方法・チャネルを通じて作品に強く惹かれる人にリーチできたりすることがある。こうしたメリットが、収益増大につながる例もある。

データの利活用についても似たようなオープン化のメリットが存在する。例えばある欧州の都市では、政府が保有する街路樹の位置データを一般に開放したところ、予想しなかつた形で利用されることになったという。利用者は花粉症などのアレルギーを持つ人々であり、通勤などの際の移動経路や住む場所を決める際にこの街路樹データを参考にするのだという。街路樹の位置データはどのような価値を持っているのか、保有者である政府には当然一定の理解はあるが、政府はデータの潜在的な価値を知り尽くしているわけではない。そこでオープン化によってさまざまな人が利用しやすくすると、政府が当初予期していなかった価値がそこから引き出されることがありうる。

ニュージーランド政府の交通庁は道路の利用データを保有していたが、2012年にこれを一般開放している。オーストラリア・ニュージーランド銀行グループ(ANZ)は、このデータを一定の形で処理すると、それが景気の状況と高い相関を示すことを見出している。トラックなどの重い車両とより軽量の自動車・バンなどを区別すると、前者は経済の現況と密に連動し、後者は6ヶ月後の景況の先行指標としてかなり有用であることが発見された。これがきっかけとなって、今日でも用いられるTruckometerと呼ばれる経済指標が開発され、ANZによって公表されるようになっている¹⁷。

このような個別例に限らず、データについて広く一般的に当てはまる傾向として、あるデータが特定の用途への利用価値があるか、例えば十分な精度で特定の判断に使えるかどうかは、実際に分析してみないとわからないことが多い。何かしらの用途に使える、価値のあるものであるのかも、必ずしも明確とは限らない¹⁸。仮にデータの保有者が様々な潜在的な用途をすべて知っていて列挙できる場合であっても、そのような用途が多く考えられる状況であれば、実際に十分に有用であるかどうかの分析をするためのコストをデータ保有者がすべて負担することは考えにくい。そこで、利用しやすい形で一般開放することによって、様々な関心と動機を持った主体がこのデータの有用性を確認し、そのような分散的な試行錯誤の結果としてデータの利用価値がより多く見いだされることになる。

ちなみに、オープン化を実施した主体もメリットが得られることがある、という点はデータのオープン化にもあえはまる。政府は時に、データに含まれる間違いの指摘を受け取ったり、第三者によるデータの分析結果から知見を得たり、データを活用するアプリが第三者によって（勝手に）開発されることで支出を抑えることができることがある。

このようなメリットがどこにどれだけ存在しているかは、事前にはわからない部分が多い。そして、当然ながらデメリットも存在している。すなわち、望ましくない利用が発生したり、データ利用が誰かに損害をもたらすようなリスクも存在している。オープン化は常に

¹⁷ New Zealand Government (undated)

¹⁸ データの利用価値を事前に知ることが難しい点については、渡邊（2018）により細かい検討がある。

メリットのみをもたらすわけではなく、また、常にメリットがデメリットを上回ることが保証されているわけでもない。ただし、人や組織は自ら資源配分をコントロールし、意思決定を自ら行うことを重視し、それを部分的に他人に委ねることには抵抗する傾向があるようだ。オープン化に過大に不安や懸念を抱く傾向がある、と言い換えてもよい。このような傾向があるため、オープン化はしばしば、事前に思っていた以上にメリットは大きく、リスクは少ない、ということになる。

行政組織や大企業などがオープン化をする場合には、そもそも失敗が許されないのであればリスクを強く警戒することが合理的である場合もあるだろう（失敗が許されないこと自体が合理的であるかはともかく）。だが、そのような場合でなくとも、デメリット側を過剰に警戒するバイアスがかかることはあるようだ。昨今の生成 AI への注目よりも前、2010 年代中盤からの深層学習による AI をめぐる議論の中では、AI がブラックボックスであることについてしばしば懸念が表明されてきた。AI に決定を委ねてしまうと、何を根拠に特定の決定がなされたのか、それが適切な決定であるのかを検証することが非常に難しい。例えば企業で人事・昇進などの決定にそのような不透明な AI を用いることは問題がある、という議論がある。このような議論をする際に見落とされることがあるのが、人間が判断をする際には、人間の判断がブラックボックスであり、しばしば偏見や差別を含んでいることが研究からも示唆されており、そのほかの面でも不適切さを含むものである、ということだ。AI は既存のデータをベースにして判断を行うことが多いため、管理職として成功するのは男性である、といった社会や職場環境の不公正な偏りに起因する男女差を誤って男女の能力差と判断してしまうリスクがある。では人間はそのような偏見から自由かと言えば、そこは人によって差があるが、誰もが完全に自由ではないだろう。また、既存のデータをベースに判断している点では人間も同じだが、そのデータに偏りがある点でも人間もまた同じである。

このように人間の判断と AI の判断には多くの共通の問題があるが、AI がより強く警戒されるとしたら、それは AI の情報処理プロセスがブラックボックスであるとか、偏りのあるデータをもとに判断しているとかいった理由であるよりも、自分達になじみのない未知の存在に対する警戒感であるかも知れない。未知の存在に対する警戒は一般的には不合理とまでは言えないが、AI の利用に関しては過剰な警戒なることがあるだろう。（もっとも、偏ったデータをもとに不透明なプロセスで情報を処理する AI に任せてしまうと差別的な判断が繰り返されるという「差別の固定化」のリスクを重く見るのであれば、それは適切な警戒かも知れない。）

本稿の主題であるデータガバナンスにより近いところでは、オープンデータについても同様の傾向を指摘することができるだろう。このような取り組みは、政府が保有する数値データや文書、画像などの情報資源を一般開放して商業利用や改変も含めて様々に利用できるようにするものであるため、その「一般」の利用者の中にどのような望ましくない主体が含まれることになるのかが懸念事項になりやすい。また、2009 年に発足した米国オバマ政権による推進と共に世界中で大きな広がりを見せた際に、欧米の政府職員などに研究のためにヒアリングをする機会がたびたびあったが、政府内で懸念が示されたり、抵抗勢力のような人々が存在する点はどこも同じであった。政府職員から出る懸念には、改変によって政府データが詐欺に使われるリスクがあるのではないか、詐欺でなくとも何か誤解を与える

ことになるのではないか、悪意を持った主体が犯罪などの計画ためにデータを用いることがあるのではないか、災害リスクや治安の悪さなどが詳細に明らかになることによって特定の地域の資産価値に悪い影響が出るのではないか、などといった点を含め、「誰にどのように使われるかわからない」ことに根差すものが少なくない。そのような利用を抑止するために、改変の禁止、犯罪行為への利用の禁止、公序良俗に反するような利用の禁止、などが検討されることがある。原理的には、メリットとデメリットのどちらが強いかについては事前に断言できないことが多いが、現在までのところ、オープンデータに制限をかけておけばよかった、というような目立った悪用例のようなものは日本にはほぼ存在しないようだ。他方、メリット側に目を向けるなら、利用しやすい形でデータを一般開放することで生まれたと思われる利用例は様々に存在しており、それは政府がデータを保有していただけでは生まれ出さなかつたと思われる恩恵を社会にもたらしている。

そもそも、詐欺のためにデータを改ざんして悪用するような主体であれば、ライセンスなどで悪用を禁止してある場合でもそれに従うとは期待しにくいだろう。その点からもこうした懸念が的外れであったり、利用条件を課すことは実効性が極めて低い対処法に留まる点にも注意が必要だろう。

もちろん、制約が常に無意味というわけではない。コンテンツの分野であれば、キャラクターなどについて一定の利用を許諾しつつも、アダルト系のコンテンツへの利用を禁止するといった例は多く見られる。こうした利用がキャラクターのイメージ戦略や、作者の心情的に歓迎できないということは十分に考えられる。禁止することで断念する利用者ばかりではないものの、アダルト系の利用を明示的に禁止しておくことで、ほかのファンや、二次創作的コンテンツの公表の場になるプラットフォームなどが、アダルト利用に基づくコンテンツを排除する動きに協力してくれることがあり、ある程度実効性はあるだろう。

以上のように、バイアスの存在は、オープン化が常に素晴らしい結果をもたらすことやトラブルをもたらさないことを保証するわけではないものの、判断を委ねる「他人」が「自分」よりも信頼できないとか、質が低いとか、危害をもたらすのではないか、という不安や懸念が過剰なものかも知れないということを意識することの重要性を意味しているだろう。

2. 5. 情報社会においてオープン化が進展する理由

情報社会においては、オープン化を採用・実施することに一定の経済的合理性がある場合がある。また、オープン化が様々な領域で進展する傾向がある。

というのも、情報社会においては、情報通信インフラや情報処理の手段が貧弱な社会に比べると、これまでよりも多くの人が手軽に情報資源を活用できるようになっており、意思決定に参加するためのコストが低くなっている。

意思決定を従来担ってきた主体の側に立ってこれを眺めると、従来と同じ意思決定コストで、従来よりも多数・多様な人の参加（インプット）を得やすい、ということになる。そして、常にではないが、多数・多様な人の参加は意思決定の質を高める可能性がある。そのような領域では、オープン化は効率性改善の手段（費用対効果の向上の手段）となる。ソフトウェア開発におけるオープンソースや、百科事典の編纂におけるウィキペディアのような参加型のアプローチはその典型と言える。動画サイトを運営することを考えても同様の事情がある。運営主体として、動画自体を自主制作するか、ライセンス料を払って購入する

か、投稿サイトにして様々な参加者に開放するか、というような選択を考えても、スマートフォンに内蔵されたビデオカメラやその編集ソフトの性能が上がるにつれ、投稿型サイトにすることの費用対効果が向上していると言える。コンテンツやライフスタイルなど一部の領域では、こうした技術環境の変化と相まって、インフルエンサーやコンテンツクリエイターなどに興味を持つ人が増えていること、そうした人々が金銭的な収入を得る手段が発達しつつあること、などもオープン化の加速要因となると言えるだろう¹⁹。

情報資源を提供し、活用を促したい主体から見ても、同様の事情が成立する。ある映像作品について評判を高めたければ、広告を制作してそれをマスメディアで流すこともできるが、ソーシャルメディア上の口コミ共有を促したり、トレーラー映像を特定のファン層に提供して自由に切り抜き映像の作成を促したり、そのほかのリミックスを促すことなどもできる。広告とマスメディアは自らコントロールできる度合いが高く、他人の参加や利用に委ねる後者のようなアプローチはコントロールできる度合いが低い。多くの人が感想の発信ができるようになっていること、多くの人が映像のリミックスが簡単にできるようになっていること、こうした人々の発信する情報が、SNSなどを通じて似た嗜好を持つ人に届きやすいこと、などから、後者の費用対効果は情報社会においては向上している、と言える。その分だけ、こうしたオープンな方式を採用することの合理性は高まっており、従来よりもよりオープンな方式を採用する事例は増えていくと想定される²⁰。

より単純化して言えば、このようにも言えるだろう。情報社会は、通信ネットワークや情報処理技術の高度化・普及によって、人と人がつながりやすくなり、情報の処理をしやすくなっている。そこで、より多数・多様な人が参加することでメリットが出る意思決定プロセスや、多数・多様な人が使うことで価値がでる情報資源は、開放が進むことになる。デメリットがあるため、常にすべてのオープン化についてメリットがデメリットを上回るとか、オープン化は進めれば進めるほどよいということではないことは既に述べたとおりだが、「プラットフォーム」「ネットワーク」と呼ばれるような性質を持つビジネスはしばしばこのようなオープン化をうまく活用しており、こうしたビジネスが様々な領域で大きな影響力を発揮しつつあることからも、情報社会におけるオープン化の合理性が伺える。

2. 6. 生成 AI の学習データとオープン性

このような観点に立つと、学習データのかなり自由な利用が可能で、取引コストや取得コストが少なくとも著作権に関しては低く抑えられていることについては、そのデータが持つ価値を引き出す後押しをすることになると評価できる。現に生成 AI の開発は国内でもベンチャー企業、大手企業、大学などの研究機関といった多様な主体により行われている²¹。学習データについては、そもそも提供者側がまとまった形で存在しておらず、クローラーなどを使って収集されている場合も多いようだが、特定のデータ利用者に対して個別に許諾

¹⁹ こうした動向を整理したものとして、例えば、五十嵐、ほか（2023, 2024）。

²⁰ 例えばファンによる n 次創作が近年の推し活を急速に拡大させたとする指摘はこのような考え方と共に通していると言えるだろう。（三菱 UFJ リサーチ＆コンサルティング, 2024）。

²¹ 総務省（2024）。特に第 I 部第 4 章第 1 節。サイバーエージェント（2023）のように、ウェブ上で公開されているデータからの学習で LLM を開発したことを公開している例もある。

を必要とする場合に比べ、かなり自由な利用を可能にしている現在の状態は、様々な創意工夫や模索が可能になる点で、AI開発の多様なアプローチを可能にする利点があるようと思われる。画像系であれテキスト系であれ、モデルを開発にするにあたってはデータを様々な基準で収集・取捨選択・加工している事が既存の論文などからは伺える。特定少數のデータセットからの選択のみであったり、そのデータセットの加工についても許諾が必要になるといった状況であったりするならば、こうした開発者側の多様な知見・着想・仮説などを開発に取り入れることは難しくなるだろう。

もっとも、より優れた状況として、単にウェブ上で様々な情報が公開されているだけでなく、収集が簡単になっているとか、それらについて詳細なメタデータが存在しており、取捨選択にあたってそのメタデータを参照することが可能である、といった状態になつていれば開発にとってより大きなメリットがある、というような可能性は考えられるだろう。まとめたデータセットは収集・提供されている例があるが、特にメタデータについては様々な限界があるようと思われる。ただし、そもそもウェブの世界は特に誰かに承認を得ることなく誰でも自由に思い思いのコンテンツを公開することができるというオープン性（参加の自由さ）に大きなメリットがある。ウェブ上のコンテンツが今日のように多種多様であること、大量であることは、メタデータ付与のような技術的な条件が課されていなかつたことのお陰であると言える面もあるだろう。

その際に、利用資格などに特に制限がないため国内でも様々な主体がAIの開発に取り組めるようになっていることや、料金が課されていないために、資本力が乏しい大学のような主体であっても取り組めていることは、オープン化の成果と言うこともできるように思われる²²。多くの主体が開発に取り組めることで、より多様な技術開発が行われる可能性もあるだろう。とりわけ大学が開発に取り組めている結果、AIに関する研究情報が公開されやすくなることにもメリットがあるだろう。（もっとも、生成AIの開発にあたっては、日本の著作権法だけを考慮すればよいとは限らず、少なくとも学習用データの収集・加工やAIモデルの開発などに用いるサーバーの所在地などによって適用される法が異なって来る部分もあると思われる。）

もうひとつ、このような学習データの取引コストの低さは、利用にあたって対価を払う必要がない（例えば補償金制度のようなものがない）ことと相まって、生成AIが提供される際の価格を下げる効果もあると思われる。これは生成AIのようにかなり汎用性の高い技術に関しては重要なメリットをもたらすことが期待できるだろう。というのも、多くの人が使えるような価格で汎用技術が提供されれば、それをどのような用途に使えるかについて多様な人が多様な発想や知識に基づいて試すことができるためだ。そのような形で多くの、多様な人が用途の探索に加わることができると、汎用技術がどのような用途に特に向いているのか、どのような用途には改良が必要なのか、といった学習が社会全体として見ると進む。これは一般的には、社会が生成AIのような汎用技術からより多くの恩恵をより早く引き出すためには有効なアプローチになる。また、そのような学習が進むことで、AIの開発やカスタマイズに関わる側も、市場の拡大がしやすくなるだろう。ここから、学習データ

²² 例えば英国ではデータの購入予算が確保できず参入を断念するケースがあるとされる。
Intellectual Property Office (2022).

をめぐる現在の状況は、生成 AI の開発と利用双方にメリットがあり、生成 AI の普及を後押しし、社会が恩恵を受けるために良い環境だ、ということになるように思われる。AI モデルの中には様々なレベルで関連情報が公開されており、二次的利用ができるような条件で提供できるものが見られる。テキスト系であれば Llama、画像系であれば Stable Diffusion は比較的そのような加工の余地があることで様々な利用やカスタマイズ版、関連アプリケーションなどの開発につながった顕著な例と言えるだろう。AI の領域で「オープン」と称するものが、本稿で述べた狭義のオープン性の基準を満たさないことはしばしば見られるが、著作権法の制約次第では、このような限定的な程度であってもオープンな性質を持つ AI の開発・公開があることは、さまざまなメリットをもたらす可能性がある²³。例えば研究開発の加速、その前段階的な学習、カスタマイズを通じた多様な（もともとの開発者が想定していなかったものも含む）用途などへの応用や洗練、多様なサービスでの利用、などが挙げられる。学習データのガバナンスがより厳しければ、こうしたモデルの公開も難しくなることも考えられないわけではない。すなわち、AI モデルが何らかの意味で学習データの二次的著作物に該当するということになれば、学習データの収集・加工だけでなく、開発された AI モデルの利用にあたっても多数の権利者の許諾が必要ということになり、おおよそ現在のような AI 開発・利用のシーンは成立しないだろう。あるいは学習データが非常に高額であれば、開発者はその成果を共有せずに、有料サービスなどの形で展開してコストを回収することにより強い関心が向くかも知れない。

もっとも、学習データのオープン性が望ましいのであれば、それを利用して開発される主要 AI モデルについても、より高い水準のオープン化を達成し、自由に研究・加工などができる方が望ましいのではないか、という考えもありうるだろう。AI モデルは著作物に該当しないと思われるため、著作権制度からはややはみ出すような政策であることからここでは手短に検討するにとどめるが、全ての AI モデルについて高い水準のオープン性を要求することが（たとえば規制によって義務付けることが）最適であるとも考えにくいものの、一般的にはオープンなモデルとそうでないモデルの競争が起こることは、オープンでないモデルのみの競争よりも望ましいように思われる。そのような競争は画像やテキストの生成 AI についてはある程度は存在しているように思われる。また、AI 開発に関するコスト要因を考えると、モデルのような情報資源のオープン化だけでは解決が困難な、物・人・エネルギー資源のボトルネックがあるようにも思われる。すなわち、GPU・データセンター、エンジニア、電力などの確保が課題であり、これらは情報資源のように万人に開放し、無償で自由に使えるようにすることは難しい。こうした資源を囲い込むことや、安価入手することができる立場にある主体は、開発競争においてある程度の優位性を得ることが考えられる。このような状況下では、情報資源部分（学習データや AI モデルなど）が高い水準でオープンであっても、AI 開発に参入するためのコストが高いレベルにとどまる可能性が十分にある。非情報資源についてもオープン化を達成するようなより広い範囲の実効性のある政策パッケージは、相当困難なものであるように思われる。

もう一点、オープン化とは弱い関係を持つにとどまるものだが、学習データの利用は比較的容易・安価ではあるものの、それが研究など一部の目的に限定されるような著作権制度を

²³ Global Partnership on Artificial Intelligence (2025).

想定すると、その場合には開発に興味を示す層が非営利系に限定されるなどして、開発は今日ほど広い層をひきつけることができない結果となるだろう。たとえば英国の著作権法はおおよそこのような規定ぶりになっており²⁴、より広い範囲の活動について著作権法の影響を弱めることも検討されたが実現していない。

3. 学習データ利用に関する制度の評価

3. 1. 学習データの利用に関する幅広い問い合わせ

生成AIの学習データとして既存の情報資源を利用するにあたっては、日本の著作権法上は、いくつかの重要な制約はあるものの、他人の著作物であっても許諾なしに利用することができるようになっている。これが望ましい制度であるかどうかについては様々な意見や利害対立があり、その対立を受けた対策なども多様な事例がある。

現行の制度が望ましいかどうか、また、どのような制度が望ましいかについてはかなりいろいろな、それぞれに難しい判断の上に結論が出るような事柄であるように思われる。現行の著作権制度だけを切り離して単独で結論が出しやすいようなものではなく、かなり幅広いほかの事柄と関連しあって評価するべきものになっている度合いが高い、と言ってもよい。あるいは、「本当にそうなのか、こういう観点から別の評価が正しいのではないか」といった異論の余地が非常に多方面に渡って存在している、と言ってもよいだろう。この幅広さをまずは確認してみたい。

3. 2. 経済効果と競争圧力

まず、生成AIの活用が生産性向上効果を持っていることが多くの調査から伺える。PCやインターネットがそうであるように、影響が及ぶ範囲もかなり広い範囲の産業にわたっている²⁵。そこで、様々な情報資源の学習データとしての利用を簡便にしておくことには、経済の効率性改善を促し、経済的な豊かさの維持・向上に貢献する効果がある可能性があるだろう。ことに、日本語のデータや日本文化の様式やティスト、習俗や風景などを反映した画像データ、そのほか日本に特殊な事情を反映したデータがあることは、それに対応した学習をしたAIが開発され、日本の利用者により意味のある知見や生成物を提供できる可能性

²⁴ Copyright, Designs and Patents Act 1988, Section 29A,
<https://www.legislation.gov.uk/ukpga/1988/48/section/29A>

²⁵ 例えば、OCED (2024)報告書では、20%以上のタスクが従来の半分以下の時間で遂行できるようになるような人という基準を満たすのは、OECD加盟国の全労働者のおおよそ4人に1人に上っており、将来的には70%に達するとしている。より大きな影響を受ける人に限ると、50%以上のタスクを従来の半分以下の時間で遂行できるようになる人は、OECD加盟国の全労働者の40%になるとしている。日本経済を対象として実施された類似の分析(新田, 2024)では、生成AIで代替される度合いが高い職業に就いている人は全就業者の21%、生産性向上につながる可能性が高い職業の人は18%としている。

ただし、こうした見方に反論がないわけではない。例えば、生成AI以外のAIを含めて扱い、生産性向上にともなうコスト削減の成果を足し合わせたものをベースにして推計した結果向こう10年間で米国経済が得られる恩恵はGDP成長率にして0.66%程度に留まり、それをさらに下回ると考える十分な理由があると指摘する論として、Acemoglu (2024)

を確保する上では重要だろう²⁶。生成 AI のビジネス用途を見ても、情報の検索やアイディアの考案、人間関係についての相談、文書やプログラムのドラフティング、プレゼンテーション用資料作成など実に多岐に渡る。少なくとも今のところはこうした利用がかなり安価に行える状況にある。

既に利用が飽和しているという様子もなく、AI の高性能化も続いているようだ。そこで、生成 AI は開発・普及の後押しにメリットがある段階であるように思われる。これがもうすぐ終わるという様子もあまりない。

加えて、経済的な競争上、AI 活用を遅らせることが企業間や労働者間の国際競争上不利になる可能性もあるだろう。産業構造が大きく変動して、それに対応できる企業は成長の機会を得、対応できない企業が衰退の圧力にさらされる、といったことが起こる可能性だ。これからも、日本の利用者に使いやすい AI が開発され、普及することが、日本政府の政策としては望ましいということになる。

一時期、現在は生成 AI と区別して識別系 AI と呼ばれることがあるような深層学習の応用可能性についての政策論議や研究の文脈では、AI が得意とするのは定型業務であるためクリエイティブな仕事は AI によって価値を失うことはない、といった類の意見が比較的広く受け入れられていたように思われるが²⁷、生成 AI の高性能化・普及と共にその点は修正を迫られており、むしろ文化、クリエイティビティなどに関わる産業に大きな変容が起こる可能性を指摘する意見も少なくないようと思われる²⁸。なお、このような文脈では、対人サービスや人間心理の理解などが AI にはできることとして挙げられることがあるが、それについても、気になる動向は存在している。2025 年 8 月に ChatGPT-5 がリリースされ、ChatGPT-4o が多くの利用者に使えなくなった際に、GPT-5 の性格が冷たいことが不評を買い、GPT-4o の復帰を求める声が多く出たことはその一つだ。²⁹ また、オンラインフォーラム上で寄せられる健康関係の質問に対して医師などの有資格者が投稿した実際の回答と、同じ質問に対して ChatGPT が生成した回答を、どちらがどちらであるかわからないようにして医師に評価させた研究では、共感を示している点でも、総合的な評価でも ChatGPT の方が人間よりも高く評価されている³⁰。

もっとも、AI 生成物を直接市場に提供する場合には、それが消費対象として消費者から受け入れられるかどうかについては、検討の余地があると思われる。AI 生成物であれば興味を失う、または、対価を支払うに値する商品とみなさないといった受け止め方もあり、そのような傾向を示唆する実証分析も存在する³¹。ただし、中長期的にはそのような受け止め

²⁶ Shcherbakov, et al. (2025)

²⁷ 例えば 2013 年に世界中で話題になり、AI による大量失業の議論のきっかけとなった研究では、知覚内容に基づく操作行動、創造的知性、社会的知性の 3 種が当時の AI にはできないタスクとされていた。これは機械学習やモバイルロボティクス領域の文献や、オックスフォード大のワークショップで得られた知見に基づくものと説明されていた。(Frey & Osborne, 2013).

²⁸ De Cremer, et al. (2023), Frey & Osborne (2024), 栄藤稔. (2023).

²⁹ IT Media (2025a), Freedman (2025)

³⁰ Ayers, et al. (2023).

³¹ 日米中 3 カ国でこのような傾向が存在することについてのアンケート調査結果の分析として 田中 (2023)

方、感受性も変化する可能性もある。作曲家は演奏の腕前が高いとは限らないことを指摘しつつ、自ら絵を描く技能は持たないが人の心に訴えかける絵をAIを使って生み出せる人が登場する可能性を指摘する論などもある³²。ビデオゲームなどではむしろAIを使っていることがアピールされることがあり、分野によって受容度が異なる部分もあると思われるが、需要度が他の領域より低いと思われるイラストや写真の領域でもCGやPCの活用はされており、その一環としてアルゴリズム処理自体は行われてきていることを考えるなら、受け入れられる可能性もあると考えてよいように思われる³³。仮に消費者がAI生成物を非常に低く評価する場合であっても、生成AIが生産性を向上させる点は概ね変わらないだろう。現状では生成物を直接販売・提供しないような利用法も多いが、既にある程度生産性が向上していると思われるためである。そして、その場合にも、学習データ利用を困難にするよりも、容易にしておく方が生産性の向上に寄与しやすいので、現行のような制度がよいということになるように思われる。文章を扱う生成AIについては特に、消費者向けに生成物を提供する以外の用途が（そもそも文章の提供を事業としていない産業も含め）幅広い産業に存在していることから、生産性向上から得られる恩恵は大きなままにとどまるだろう。

関連した論点として、生産性向上のほかにも、労働の質の向上や、労働時間以外の時間での充実なども考えられる。例えば日本で生成AIを現に業務に使っている人たちに仕事への影響を尋ねた2024年の大規模調査では、仕事のパフォーマンスへの影響以外に、その他の面（有給休暇、残業時間など）についても尋ねているが、いずれも改善がみられると回答した人が、悪化がみられると回答した人を上回っている（なお、仕事のパフォーマンスについては改善した人は60%を超え、低下した人は11%を超える程度という回答分布であった）³⁴。

3. 3. 経済的損失の不均一な分布

生産性が向上する場合でも、すべての人が等しく恩恵を受けるわけではなく、むしろ経済的な損失を被る人も存在する。AIの学習データ利用がかなり直接的に自社・自分の市場価値の低減につながるのではないかと思われるケースがある。俳優が自らのデータを提供す

³² 深澤（2024）；類例として、ソフトウェア情報センター（2024）の第一回会議で言及されている魯山人のエピソードを挙げることもできよう。より一般的には大きな予算と人員で取り組むような創作・制作では、直接手を動かす人と、指示出し・取捨選択などをするとの分業が発生しやすく、著名な建築家の建築であれ、アニメーション作品であれ、「作家」として一般に認知されている人が自ら手を動かして線を引いているとは限らず、作家の作風は指示出しや取捨選択を通じて發揮されている状況が既にあるとみることもできるだろうと思われる。

³³ マンガ分野で類似の見解を示しているものに、小沢（2024）がある。当該セミナーでは次のようにも述べている。「アシスタント、スクリーントーン、コピー機、3DCG、デジタル作画、これらは出た当初ずるいと言われたのですが、今ではもう当たり前に使っています。」ほかに、ソフトウェア情報センター（2024）の研究会第5回でも、AI生成物の受容度が低いのは過渡的な現象ではないかという説が検討されている。

³⁴ 天瀬、ほか（2025）；実証研究ではないが、逆に生成AIが様々な職業のスキルの格差を圧縮し、競争が激しくなるが需要が伸びない結果として所得が減少する、とする説もある。（Frey & Osborne, 2024）。

ると AI でよく似た人物を合成できるようになる、声優やイラストレーターも同様の部分がある、といったケースが典型だろうと思われる。アート、エンターテイメントなどの産業を分析した最近の研究では、作家、ライター、翻訳者、通訳者、などは生成 AI による自動化の可能性が高いとするものもある³⁵。

経済的な問題として考えるなら、一般に、ある技術の利用がもたらす経済的な利得と損失を比較して利得の方が大きいのであれば、損失を被る人に何かの形で利益の再分配を行うことが、技術の利用を止めるよりも望ましい。ここでも、学習データ利用は促し、そこから得られるメリットを大きくしつつ、再分配政策をセットで実施することが望ましいということになるだろう。

ただし、社会全体の経済的損得のみに着目するならそのような結論になるとしても、「あなたの創作物やあなたの実演データを提供してほしい、それを AI に学習させてあなたの仕事を代替したい」、と言われて納得する人は多くないだろう。拒否する手段があるなら拒否し、AI の開発・普及を遅らせる方がこうした人々にとっては経済的に合理的な選択になっている場合があると思われる。また、損得勘定以外の価値判断としても、そのような AI の開発・普及には納得しがたいと考えると思われる。実際にコンテンツ産業やクリエイター、実演家などに関連する団体で無断学習に反対する意見を表明しているところは非常に多く見つけることができる³⁶。無断学習に関連した訴訟についても、同様の観点から理解できる部分があるだろう³⁷。

³⁵ Berg, et al. (2025a). 対象としている文化・メディア諸産業は高所得国では雇用の 1.8% 弱を占めるとしており、その中でも代替されるリスクの高さは職業によってかなりばらつきがあるとの結論に至っている。日本語による解説として、労働政策研究・研修機構 (2025a)。

³⁶ 例えば生成 AI による無断学習に懸念を表明したものとして、日本美術著作権連合 (2023). 生成 AI に関する声明, 2023.08.17. https://www.jart.tokyo/jart_wp/wp-content/uploads/2023/08/5bbd651f0394f60cbdbf80279ca1b5ac.pdf; 日本雑誌協会、日本写真著作権協会、日本書籍出版協会、日本新聞協会 (2023). 生成 AI に関する共同声明. 2023.08.17.

https://www.pressnet.or.jp/statement/copyright/230817_15114.html; 日本写真家協会 (2023). 生成 AI 画像についてその考え方の提言 <https://www.jps.gr.jp/about-generated-ai-images/>; 日本脚本家連盟、日本シナリオ作家協会 (2023). 脚本家の権利と生成 AI に関する共同声明:2023 年 11 月 3 日文化の日に, 2023.11.03. <https://mitte-x-img.istsw.jp/scenario/file/kyodoseimei20231103.pdf>。音楽産業のステークホルダーからの生成 AI 批判として日本音楽著作権協会 (2023). 生成 AI と著作権の問題に関する基本的な考え方, プレスリリース, 2023.07.24,

https://www.jasrac.or.jp/information/release/23/07_3.html; 日本音楽著作権協会 (2024).

「AI に関する音楽団体協議会」の設置について: for Creators, for Artists, 2024.01.25. https://www.jasrac.or.jp/information/release/24/01_3.html。また、生成 AI 全般ではなく、特定の声優の声を無断で追加学習に用い、その成果を公開することに対する抗議として、『NOMORE 無断生成 AI』声優有志の会 (2024). Statement, <https://nomore-mudan.com/#statement>。ほかに、アニメ業界の関係者へのアンケート調査から AI 規制への意見を探ったものに日本アニメフィルム文化連盟 (2023)。音楽産業の被る 2028 年時点の経済的損失を試算したものとして、PMP Consulting (2024)。

³⁷ 米国における訴訟は数が多いが、例えば Vilendreer (2025) に比較的最近の概観があ

もっとも、著名なイラストレーターなどを想定すれば、そのような人の画風を真似て類似のイラスト画像を生成できるように追加学習モジュールを開発することや、そのようなモジュールを活用して画像生成を行うことが著作権侵害になる場合も多く出てくるだろう。というのも画風だけを模倣するに留まるのであれば侵害にならないかも知れないが、その作家に特徴的なキャラクターやモチーフまで再現した場合にはそれが侵害になりやすいと思われるためだ³⁸。そうであっても、AIの普及によって侵害が蔓延してしまうことを想定すると、個別に削除要請や訴訟を通じた賠償金の収集などを行うことは多大な労力を要する。そこで、開発・普及が進まないことが望ましいと考えることが経済的に合理的になることも考えられる。ただし、生成AIを活用して効率性を向上させることで、生成AIから恩恵を受けるようになる可能性もあるようにも思われるため、細かくは断言できない部分が多いように思われる。

非常に有名なプロフェッショナルから完全に趣味で活動しているアマチュアまで様々な人がいる中で、うまく大半の利害がまとまるということはないかも知れない。その場合には、学習データとしての利用に特に反対しない人のデータがAIに利用されることになるだろう。AIの開発・普及を遅らせる努力をする人、気にしない人、AIを活用する側に回ろうとする人などにより多様な対応が出てくることも十分考えられるだろう。そのような状況になり、結果としてAIが普及し、それによって所得が減り、転職などが必要になった場合、再分配政策・支援策もやや複雑になることが考えられる。生成AIの開発・普及に反対し、自らも利用しなかった人々が最も経済的に損失を被ったとして、こうした人々に対して最も手厚く支援をするべきなのか、それともより早い段階で業態変更や他業界への転職などの形で適応の努力をした人と同じ支援をするべきなのかは自明ではないように思われるためだ。

以上のような政策の詳細を決める上での困難があるとしても、学習データをめぐる取引コストを引き上げて、生成AIの開発・普及を遅らせた方がよい、ということにはならないと思われる。もっとも、現行の著作権制度にもある、「著作権者の利益を不当に害することになる場合には」無断利用を認めない、という規定の解釈次第では、このような、生成AIによって比較的直接代替される可能性が高い作品などを学習データとして用いる場合には許諾が必要である、とする考え方もあり得るだろう³⁹。

る。他に Weisenberger, et al. (not dated).

報道記事では、Dent, S. (2024), Barcott, B. (2024). Lutkevich & Heaton (2025).

学習に用いられたデータが生成画像にそのまままたはかなり類似した形で出てくることがある点については、例えばソフトウェア情報センター(2024)の第一回会合での議論を参照。

³⁸ 学説上は、生成AIを含め、情報解析に該当する行為はこのような侵害にはあたらないとする解釈もある。この点についての批判的検討として、奥村弘司(2025)なお、そのような説を探った場合でも開発は侵害に当たらないとしても、画像生成については侵害に当たると考えられるように思われる。

³⁹ 文化審議会著作権分科会法制度小委員会(2024) p.23-24には、この点について興味深い記載がある。無断利用は認められるとの解釈を採用しつつ、そうでないとする意見もあることを記し、著作権法以外の法に照らした問題がある場合についても言及している。奥村(2025)にこの点についての議論の概観・検討がある。

損失にだけ着目するなら、一部の学習データの直接の「提供者」に当たる人（提供したいという意思を持っているかは別として）以外にも損失を被る可能性のある人々は考えられる。そして、このような損失を被る人も社会に均一に分布しているわけではない。生成 AI の雇用・経済へのインパクト分析を行った研究の中には、AI を使いこなせない中高年の労働者⁴⁰、熟練度が低い仕事にしか就けない若年層⁴¹、AI の影響を受けて代替が起こりやすい職に就いている率が相対的に高い女性⁴²、などが経済的に損失を被るのではないかと示唆する研究なども存在している。また、AI によって影響を受けるセクターは都市部や高スキルの仕事に多いなど傾向があり、その内、一部は AI によって代替されるが、一部は AI によって補強されるという見方もある。この場合は、解像度を上げなければ多くの恩恵を得る人と、多くの被害を受ける人が重なる、というようなことにもなるだろう。

生成 AI がどの程度の損害を発生させるかについて検討する際には、失業の可能性が取り上げられることが多い。だが、より広い範囲の人々が業務のやり方を変える必要が出る、部分的には生成 AI に代替される、などの影響を受けることになると思われる。生成 AI によって代替される可能性が高いとされる職に就いているのは全世界の労働者のおおよそ 3.3%だが、何らかの形で影響を受ける労働者は 4 人に 1 人、とする分析もある⁴³。

失業や配置転換、業務内容の変更などを迫られる人々は、他人が提供した学習データに基づいて構築された AI モデルによって損失を被っている人々であると見ることもできるだろう。支援策はこうした人々も対象にする必要があろう。リスクリリング支援などはまさにこれに適した政策であろう。ただし、リスクリリングがどの程度成功を収めることになるかについては注意が必要かも知れない。うまくリスクリリングができる層とそうでない層が出る可能性や、こうした不均一な結果が既存の経済格差を拡大させる可能性なども考慮に値すると思われる⁴⁴。

こと昨今の生成 AI に関しては、米国の巨大テック企業による大規模投資が開発をリードしているとの見方に立ち、AI 利用がそもそも日本経済を利することになるのかどうかを疑

⁴⁰ Cazzaniga, et al. (2024).

⁴¹ ソフトウェア開発やカスタマーサポートなどのセクターでは若年層の雇用が減っていることや若年層の雇用が全体的に伸び悩んでいることを大規模給与データの分析から示唆した研究として、Brynjolfsson, et al. (2025)。

⁴² Gmyrek, et al. (2023). ただし、この研究では女性は恩恵を受けやすい職に就いている割合も相対的に男性より高いとの結果も示している。;AI に影響を受けやすい職に就いている割合は大差がないものの、要求されるスキルの水準がより低いため、結果としてより代替されやすい可能性があると指摘したものに Council of Economic Advisors (2024); 内閣府 政策統括官（経済財政分析担当）(2024)。

⁴³ Cazzaniga, et al. (2024) では影響を受けるのは 40% 程度の労働者であり、20% 程度は生成 AI が人間を補完する形に、20% 程度は生成 AI が人間を代替する形になるとしている。他に、Berg, et al (2025b) では 4 人に一人の労働者が何らかの形で生成 AI に影響を受け、3.3% (つまり 30 人に一人) 程度の労働者は特に大きな影響を受ける (4 段階分類で最も影響の大きい段階に属す) と推測している。ただし、この影響はほとんどの場合失業ではなく、業務内容の変化などの形をとるとしている。日本語での解説として 労働政策研究・研修機構(2025b)

⁴⁴ Cazzaniga, et al. (2024).

問視する向きもある⁴⁵。これはいわば、日本全体が損失を受けるとか、恩恵をあまり受けられないと見る見方だと言えるだろう。昨今の「デジタル貿易赤字」論もそのような懸念を伴うことがあるようだ。経済的な問題だけを考えるならば、デジタル貿易赤字だけに着目する理由はそれほどなく、むしろ貿易収支全体を考える方がよいだろう。特にデジタル系の財やサービスを輸入していることで国内産業の生産性が高まっているなら、それを踏まえた判断が必要になるだろう。動画編集ソフトであれ、文書作成ソフトであれ、PC や携帯端末であれ、AI であれ、日本企業によって開発・提供される代替品や、オープンソースなど無償で提供されているものと比較して海外のものが優れているなら、そちらを使う方が日本経済にとってもよいし、貿易収支にもよい、という可能性は十分にある。デジタル貿易赤字は貿易収支上小さい規模ではないため、通貨レートへの影響なども考えると政府としては放置することが最適とは限らないが、国内産業の成長を支援することは有益であっても、海外製のデジタルサービスへ関税をかけるとか、利用を阻止するといった政策は得策ではないだろう⁴⁶。

3. 4. 恩恵の不均一な分布

損害を被る人ではなく、利益を受け取る人の側に着目した議論も、多く存在する。AI 開発を主導している米国の大手企業やその従業員を主要な受益者と考え、このような人々を利することを問題視する意見もある。

直前の議論と重複するが、注目されやすいのは米国の巨大テック企業であろう。OpenAI や Anthropic が典型的な常利組織になっていないことを脇に置くとすると、Google、Meta、Microsoft などの米大手テック企業は生成 AI のモデル開発においても主要プレイヤーとして活躍している感がある。もっとも、こうした企業の影響力に対する警戒や政策対応は、日米欧を含む様々な国や地域で起きている事である。こうした企業が AI の普及によって大きな利益を得ることになるとして、また、生成 AI に関する市場の大きなシェアを獲得することになるとして、そこから発生する問題は第一義的には競争法（日本では独占禁止法）が担当する領域であろう。企業規模や市場シェア自体ではなく、その影響力を経済に損害を与える形で行使することが問題にされる。競争法をめぐる学術的な議論は近年盛んであり、いわゆるプラットフォーム規制のあり方も含め、今後の研究の進展によってより厳しい規制が標準的になっていく可能性なども考えられるだろう。また、欧州ではデジタル市場法や AI 法などを通じた規制があり、日本ではいわゆるデジタルプラットフォーム取引透明化法やスマートウェア競争促進法を通じて独禁法以外の形で競争環境の維持・確保を図る制度が近年できている。このような制度があれば問題が起きないかというとそのような保証

⁴⁵ 例えば著作権法を中心とする議論の中でも、類似の意見をソフトウェア情報センター（2024）の会議（第1回）に見ることができる。

⁴⁶ デジタル貿易赤字単体で評価すべきでないという意見として、西角、綿谷（2024）；為替レートへの影響は要注視、デジタル産業が国内に十分育っていないことを問題としつつ、デジタル化の経済メリットも指摘する論として、岳（2024）；デジタル貿易赤字の問題を深刻なものと受け止めつつ、一方で貿易収支改善のためのカギとなる産業の海外展開支援、プラットフォーム型事業支援などを重視し、海外の効率のよいデジタルインフラの利用は継続することをよい選択とする論として、津田、ほか（2025）

はないが、起こりにくく、問題が発生した場合に早期に発見や修正しやすいと期待することはできるだろう。また、競争法は市場競争が阻害され、被害が発生した後にそれを調査し、是正する形で作用することが多くある（合併などの審査は除く）。これが訴訟を通じて決着されることもあることもあり、問題の解消には長い時間がかかる可能性がある。その点、これら 2 つの法はより短期的な対応・解決に有益である可能性もあるようと思われる。といふのも、デジタルプラットフォーム取引透明化法はオンラインモール、アリストア、デジタル広告など特定の領域のプラットフォーム事業者に毎年の報告・自己評価などを義務付け、行政はそこに問題があれば公取委に対応を要請することもできる仕組みとなっている。また、スマホソフトウェア競争促進法は、特定の行為群を規制することによっていわば事前規制的に競争の阻害を防止する法となっているためである。

現在の生成 AI に関する市場はアプリケーションなども含めて考えると既に複雑な構造になっていて簡単に形容することが難しいが、公正取引委員会の採用した 3 層分類⁴⁷に沿って整理するとすると、本稿で主に検討している生成 AI のモデル開発に着目するなら、いわゆるネットワーク効果やスイッチングコストの高い状況ではなく、複数の AI モデルの使い分けや、同時並行的な利用が容易であるように見えることや、価格や品質の激しい競争があるように見えることなどから、直ちに介入が必要な状況とは言えないよう見える。ただし、投資規模が大きいほど性能が高まるといふ「スケーリング則」の傾向が今後も続くのであれば、その一つの帰結として、特定の事業者が多くの収益を上げられる構造ができるとそれを再度開発投資に回して更に優れた AI の開発で競合他社に差をつける、というようなことができるようになるのかどうか（いわゆる自然独占傾向が存在するのか）は、注意が必要ということになろう。また、計算資源などインフラについては現状ある程度の垂直統合や寡占性の高さが認められる、アプリケーション側も OS などに寡占が認められる、と密接に関連する市場で支配的な高いシェアを持つ事業者がそれを活用して生成 AI のモデル開発などの競争を阻害するような展開があるかは注視することになるだろう⁴⁸。だが、現状では政府が介入するほどの事態にはなっていないよう思われる。

ちなみに、AI 開発を主導しているのが米国の巨大テック企業のみかというと、違う見方も出来るだろう。OpenAI 社や Anthropic 社をどう評価するかという先に触れた点以外にも検討すべき点はある。LLM と呼ばれる領域では米国勢がリードしていることについては異論の余地はないと思われるが、DeepSeek のような中国企業の開発する AI をどう評価するかについては検討の余地があるだろう。また Mistral 社はフランスの企業である。本稿執筆中には、スイスの大学連合が開発したオープンソースの LLM の報道もあったところだ。その性能は最先端モデルほどではないようだが⁴⁹、公表されている特徴には注目すべき点があるよう見える⁵⁰。画像生成の分野では米国企業による AI の活躍もあるが（初期であれば OpenAI 社の DALL-E や Midjourney 社の同名の AI） Stability AI（英、Stable

⁴⁷ ほかに、クラウドの計算資源とは区別してそこに配置されることになる高性能チップの製造能力にも着目した分析として Artificial Analysis (2025)

⁴⁸ 比較的最近の現状分析として、王 (2025)および公正取引委員会 (2025)。いずれの分析でも、現状のところは政府介入を必要とする問題は発見されていない。

⁴⁹ Dunn (2025).

⁵⁰ Moon (2025).

Diffusion シリーズ) や Black Forest Lab (独、FLUX シリーズ) といった欧州地域を拠点とした企業の活躍も目立つ。中国企業の開発した画像生成 AI でも性能が高いものとしてアリババ社の Qwen-Image や HiDream 社の同名のモデル、ByteDance 社の Seed などがある⁵¹。

このような状況を踏まえると、現状では経済的に問題のあるような形で（公正な競争が阻害され、それによって経済が不効率になり、消費者や社会が損失を受けるような形で）の恩恵の偏在はない、ということになろう。そうであれば、大きな恩恵を受けている企業を直接ターゲットにするのではなく、一般的な再分配政策によって損失や恩恵の偏在を是正していくことが重要ということになるだろう。ということは、学習用データの利用しやすさについて、特定の開発者に対しては利用を困難にする、それ以外の開発者を優遇する、といったような政策もまた不要で、AI から受ける恩恵を大きくするような現行の制度で現状の所はよい、ということになると思われる。

なお、恩恵の不均一な分布の観点からは、もうひとつ異なる可能性が懸念事項として挙がることがある。それは AI の普及率の格差や、そこからの帰結としての経済的な恩恵の格差だ。都市部の労働者や、高スキル職の方が生成 AI を活用しがちであることなどから、これが地域格差、所得格差などの拡大につながる可能性を指摘する研究も存在している⁵²。著作権法や学習データのガバナンスとは関係は薄いが、政策課題としてはこのような AI 利用者側の格差への取り組みが当面重要な可能性も十分考えられるだろう。

3. 5. 経済安全保障、経済的搾取の倫理的問題

以上はいずれも経済価値について、オーソドックスな経済学が扱うような点に着目した整理だが、学習データの利用を容易にするべきかどうかについては、それ以外の論点も多く存在している。経済的な価値に近いところで言えば、経済安全保障の観点がある。すなわち、AI は汎用的な国のインフラであり、それを外国の企業に依存した形で調達することには安全保障上の問題がある、という考え方だ。中国企業によって開発・提供されている AI への警戒や注意喚起は日本でも比較的されてきていると筆者は考えているが、米国企業によるものであっても、昨今の関税をめぐるトランプ政権の動きや、第一次トランプ政権時の国際協調の枠組みに対する対応ぶりなどを見ると、米国をどの程度安定的なパートナー国と考えるべきかについては多様な意見があり得るところだろう。ただし、現在ある高性能生成 AI はいずれも開発・利用に膨大なデータ処理を伴うものであり、仮に AI だけを国産化できたところで、クラウド基盤が安価に用意できなければ（米国企業が提供するより安価なクラウドの上で開発・提供される AI を利用する場合に比べて）効率性改善効果が落ちることになり、市場競争や国際競争で優位に立てる可能性も減ることになろう。単に生成 AI モデルの開発だけでなく、それに必要なインフラなども整備することが望ましい、ということになる

⁵¹ 主要な AI がどれであるかを決めるることは容易ではないが、米国の報道機関によるレビューの例としては Chedraoui (2025), Griffith (2025), Ortiz (2025).

ほかに、ユーザー参加型のブラインドテストをとりいたるものとして LM Arena が参考になる。LM Arena についての紹介としては、Wright (2025).

⁵² Organization for Economic Co-operation and Development (2024). 日本語による解説として労働政策研究・研修機構(2025c).

が、これもまた、学習データの利用しやすさを落とす理由にはならないと思われる。

ただし、この経済安全保障の観点からは、市場競争が現在よりも低下し、寡占化が進んだ場合などには、特定の海外事業者への依存度を低くすることが特に重要になり、場合によつては、生産性向上を犠牲にしてでも、優れた AI の利用を抑制するといった政策が望ましいということになる可能性も考えられなくはない。

AI 開発・提供企業の国籍を問わず、経済的な搾取があることが問題だという意見も存在する。仮に経済の効率化に貢献するとしても、仮に米国の巨大テック企業だけがここに関与しているわけではないとしても、自分が生み出した情報資源が学習データとして利用されることによって自分の経済価値が大きく減ってしまう可能性のある人々にとっては、そのような事態をもたらし、大きなメリットを得るのが巨大な資本・投資に支えられた企業の高所得者であり、被害者側は個人事業主などの経済的に不安定な者を多く含む、といった構図は考えられよう。このような弱肉強食は市場経済の効率化に貢献するとしても倫理的に問題がある、という批判も考えられる。この倫理的な観点は、これを理由に生成 AI のもたらすメリットを断念することが妥当というほど重いものではないのではないか、という風にも考えるが、様々な考え方があるところで、倫理的な問題がある以上生成 AI をボイコットするべきであり、それによって幅広い人の利便性の向上をあきらめることになんでもそれが望ましい、という考え方もあるように思われる。現在のところ、コンテンツ産業、クリエイターや実演家などの団体から出されている声明などでは、ボイコットが適切とするような強い意見は見当たらない。むしろ、現行の著作権制度には懸念を表明しつつも、生成 AI を適切に使うことで得られるメリットに理解を示すものが多い。だが、ネット上ではより強い反対意見やボイコット運動的な動きを見つけることも出来る。そのような観点に立てば、生成 AI の開発を容易にするような現行の著作権制度も問題があるということになろう。

3. 6. 文化的損失

もうひとつ、本稿で明確な解決を示すことが難しいと考えている点がある。それは文化の大きな変容、担い手の喪失、それに伴って発生しうる既存の文化の衰退の是非だ。このような変化が起こるかどうか、確実なことが言えるわけではないが、大きな変化が起こるとする見方は珍しいものではない。既に言及したコンテンツ産業、クリエイターや実演家などの団体にも、研究者にもそうした見方をする例は見られる。

一面では、結果として文化により広い範囲の人が参加することができるようになる点は、よい面もあると思われる。例えば料理やスポーツのように著作権があまり広く及ばない領域で成立している文化を考えるならば、多くの人が参加し、技巧のレベルはさまざまながらも、参加を楽しみ、創意工夫や研鑽にはげみ、個性を發揮し、その成果物やパフォーマンスが受け手などの心を動かす、といったことが起こっていると言えるだろう。これらの領域のごく一部には、プロフェッショナルの高度な領域（プロスポーツやオリンピック、あるいは高級レストランなど）も含まれるが、裾野は非常に広く、文化の価値はその高度な領域の達成度だけで評価されるわけではない。画像生成 AI がもたらすかも知れない文化の変容として、「作風」のように著作権ではあまり保護されることがなかった側面の活用が非常に容易になる結果、人々が様々な画風を楽しみ、所望の画像を簡単に生成することができるようになり、絵画やイラスト、写真などの領域が料理やスポーツに近くなつたとして、そのような

文化にも価値がある、ということは言えるだろう。

他方、従来であれば独自の作風を強みとして生計を立てていたようなクリエイターが大きく所得を落とすことになってしまうかも知れない。既存の文化をデータとして学習対象にした場合に、それ以降はそれに匹敵する新たな学習データを生み出す者が激減する、というような変容が起こるとなったら、文化の発展に寄与するという著作権法の目的に照らしてもよいことであるのかどうか、判断の難しいところである。いわゆるプロフェッショナルだけではなく、アマチュアであっても、創作の技法を学び、時間をかけて一枚の作品を創作してきたような作り手も、生成AIの生み出す画像の物量の前には埋没してしまい、受け取れる収入やアテンションの量も低下し、創作の意欲が減退するということも起こるのではないだろうか。受け手側の一部は、(ほかの領域でそうであるように)「手作り」であることを重んじてくれることから、完全に受け手がいなくなるということはないかも知れないが、大量生産品の前に霞んでしまう手芸品のようになってしまうかも知れない。そのような形で既存の作り手が減ることはよいことだろうか。

料理やスポーツとのアナロジーで考えるならば、著作権法の保護がない領域であっても、新しいスタイルの考案や新しいプレイ方法の考案などは起こるので、新しいスタイルなどの興隆については、そこまで深刻に案ずることではない、とも思われる。しかも、そのような新しい流行は、料理であればレシピサイトであったり、ソーシャルメディアであったりといった、著名性や大資本と縁が薄いところからも現れて、広まっていく。そうであれば、現在とは大きく変容するかも知れず、現在の文化の良さは失われるところもあるかも知れないが、別の意味でよいところのある文化が生まれることになる、と言えるかも知れない。例えば、追加学習モジュール(LoRAなど)が多く共有されているウェブサイト Civit.aiなどをみると、特定の作家、作品、キャラクターなどを再現するための追加学習モジュールではなく、様々な画風の実現のための追加学習モジュールなども多く開発・公開されている。また、生成する画像内のあらゆるものをガラス細工のようにしてしまうとか、半導体基板でできているかのようにしてしまうとか、ブルーベリーでできているかのようにしてしまう、といったような生成AIの強みを活かした作風の試みなども存在しており、既存のクリエイターが収益やアテンションを得られにくくなると新しい創造活動と呼べるもののが枯渇する、というシナリオとは違う可能性も感じられる。

さまざまな文化的潮流は時代と共に変容し、隆盛・衰退もして来ており、衰退した文化については個々に惜しみ、懐かしむこともできるが、それは後に出てきた文化が劣っていることを意味しないと思われる。非常に高度な技能と投資(練習なども含め)に裏打ちされた表現などが生み出せる文化として、例えばクラシック音楽が思い当たる。クラシック音楽は現在ではポップスなどのほかの音楽ほどの市場規模を持っていない。ある意味ではクラシック音楽は優れた文化だとも言えるだろうが、ではその隆盛を維持するために音楽文化の変遷を止めるべきであったかというと、そうでもないように思われる。では生成AIがもたらすかも知れない文化変容についても同様に考えられるであろうか。

念のために付け加えておくと、著作権法は既存の文化を守り続けることを目的にしているわけではないし、そのような方針で運用してしまっては既得権益保護と大差ないものになる危険も高すぎるだろう。例えばストック写真の世界は非常に大きく変容し、デジタルカメラやインターネットの普及と共にいわゆるプロの写真家ではない幅広い層が参入できる

「マイクロストックフォト」と呼ばれるようなサイト・事業が立ち上がり⁵³、プロの写真家ほどの質ではない写真を多く提供するようになり、単価が数十万円から数十円にまで下落するという変容が起きたとされる。このようなマイクロストックフォトのサービスの隆盛と共に、当然ながら写真家として生計を立てて来た者が生計を立てられなくなり、多くの人が小銭を稼ぐようになるようになり、単価が下がったおかげで利用者は増える、といったことが起こったようだ⁵⁴。結婚式などの記念写真についても、プロの写真家が雇われることが減るにつれ、プロの写真家の所得を減じることになった。ここでもまた、写真の質は落ちているとする意見がある⁵⁵。このような作品の質の低下を伴うような変化が著作権法改正で止めるべき事柄として大きなアジェンダになったことはなかった。もっとも、生成AIについていえば、学習データの利用をどの程度容易にするかは著作権法によって大きく左右される事柄であり、ストックフォトのように経済的な競争や技術環境の変化を背景にした文化変容とは事情が違うと考えることはできるだろう。

3. 7. 長期的な負の影響: 軍事利用

AI の影響力は、国際安全保障の文脈では、核兵器に喩えられることもある⁵⁶。AI の軍事利用は既に進んでいるが、今後非常に強力な軍事力につながるものだとすると、これを自由な開発競争に任せておいてよいのか、という疑問を呈することもできるだろう。また、自由競争の結果として今日の核兵器の保有状況とそれがもたらす大規模な殺傷リスクに似た状況が生じてしまう可能性を考えるなら、AI の拡散を抑止することやガバナンス上の協調路線を確立することが非常に重要なのではないか、という考え方もあり得るだろう。冷戦時代に米ソ対立が核の開発・軍拡競争を招いたように、近年の米中の緊張関係が AI 兵器の開発・軍拡競争を招くとしたらどうだろうか。例えば核拡散防止条約のような国際条約や、反核運動のような市民運動と類似の対抗運動が必要になるだろうか。核兵器のアナロジーは使っていないものの、第二期特朗普政権が打ち出した米国の AI 行動計画は、グローバルなパワーバランスに AI が変動させる可能性などを念頭に米国の支配的地位 (dominance) 獲得を重視するというものになっている⁵⁷。

軍事力に近いところで生成 AI について話題になりやすいのはバイオテロや生物兵器の開発や製造がより広い主体に可能になるリスクであり、特に生成 AI がクラウド上で運用される必要がなくなるにつれて、あるいはオープンウェイトモデルなどの形で誰でも利用できるものとして配布されるようになるにつれて、生成 AI モデル自体の自己検閲性能などリス

⁵³ 作り手（撮影者）としての参加者の数は増えるかも知れないが、大手マイクロストックフォトサイトの参加者が多様とは限らないことを示唆する研究がある。Brabham (2008).

⁵⁴ Taub (2007), Clifford (2010).

⁵⁵ Jeffries (2013).

⁵⁶ AI と核の比喩は様々なバリエーションがあり、軍事的な文脈にも限られない。原子力技術に喩えられる場合もある（AI 同様に民生・軍事両面の利用が可能な技術であるなどの理由による）。多面的に議論したものとして例えば Matthews (2023), Roehrlrich (2025), Klyman & Piliero (2024). 他のアナロジーと合わせて概観したものとして Stern (2023).

⁵⁷ The White House (2025). より直接的に、昨今の米中の AI 競争を論じる議会の公聴会の文脈で議員が米国とヒットラー政権下のドイツの原子爆弾開発競争をアナロジーとして用いる発言もある。（Algorithms and Authoritarians: Why U.S. AI Must Lead, 2025）

ク管理が重要になる可能性はあるかも知れない⁵⁸。

このようなリスク管理が機能しないとしたら、現行の著作権制度は、むしろAI開発の進展を加速させる悪い見本だという考える向きもあるだろう。だが、核武装についての意見が必ずしも一致しないように、強力な軍事力になるのであれば日本でも開発を促進するべきだ、という意見もあるだろう。

AIの軍事利用を抑止すべきという立場からは、特定の領域の情報資源はむしろ利用できない方がよいのではないか、という考え方もあるだろう。もっとも、日本の著作権制度でできることは非常に限られているだろう。

長期的には人類全体にとってネガティブな影響も大きい事象になるのかは自明ではない部分が残っているように思われ、部分的にはAIをめぐるガバナンスのあり方にもよると思われる。学習データを日本として制限することは、単独ではほとんど意味がないと思われるが、仮にこうしたニーズがあるとすれば、政府のとるべき路線は、当面は日本も有力なAI開発・利用国でありつつ、国際的な議論の場などを通じてガバナンスための働きかけをしていくことにあるように思われる。これはいわゆる広島プロセスや、フレンズグループとの連携・協議プロセスを通じて日本が現に採用してきた路線とそう遠くないものであると言える。ただし、そもそも広島プロセスはG7に端を発しておりロシアを排除していること、中国はこうした国際協調の場から完全に不在ではないが、民主主義国と歩調を合わせて合意形成をしているわけではないこと、などを考えると、現行の路線では不十分という意見もありえるところだろう。

もう一方では米国が主導し、日本も協力している先端半導体やその製造技術・機器に関する輸出規制がある。これは製造に必須な要素の拡散を制限することで、高度なAIの開発に歯止めをかけるものと言えるだろう。（核兵器分野におけるプルトニウムやウラン²³⁵、それらを取り扱う施設の規制に喻えて考えられるところもあるだろう。）高性能な生成AIの開発・運用が夥しい量の計算資源を必要とする状況が続くことになるのかどうかが自明ではない点は、控えめに行ってもかなり気がかりであることは確かだろう。また、今後も高度なAIの開発競争が続き、その文脈では計算資源の莫大な投入が必要になる状況が続くとしても、十分に軍用に耐えるAIがより小規模の計算資源で開発・運用できるようになる可能性もあると思われる。

AIに関する長期的リスクの議論の中には（それが生成AIの一種として実現することになるかはともかく）人間の知性をはるかにしのぐような知性による人類の支配や、人間による制御からの離脱、人間に対する加害、といった更に大きなリスク（人類存亡系のリスク）についての議論も含まれる⁵⁹。また、逆に直近で具体的に報告されている悪用例から見ると、サイバー攻撃やオンライン詐欺、世論誘導のための影響力工作などで、国家が関与していると想定されるものも散見される⁶⁰。

本稿ではこうした国際安全保障や、より広く安全保障に関わるリスク（国内のテロなど）

⁵⁸ 例えばウイルス兵器開発について Service (2023)、化学兵器製造については、Boiko et al. (2023)。核兵器製造の技術的知識がより広い主体に広まるリスクについてはAllison & Herzog (2025)。

⁵⁹ 例えばテグマーク (2019)、ボストロム (2017)。

⁶⁰ Anthropic (2025), OpenAI (2025b)。

については、現状起きているような取り組みが当面は正しい方向であるという前提に立っている。すなわち、開発競争が現にある以上、自国だけが開発を遅らせることにはあまり意味がないが、リスク低減のための国際協調や、特に懸念がある主体の開発に対する牽制を同時に進める、といった路線だ。では開発競争が軍事利用の高度化も進めてしまい、そのリスク抑制策が失敗するしたらどうするのか、という問い合わせ残るように思われるが、その場合でも著作権法で解決・緩和することはできない、と言えるように思われる。日本は生成 AI 開発で最先端を行っているわけではないため、日本で開発された技術が AI の高度化に拍車をかけ、軍事利用の被害を拡大させる、ということが起こりにくく、米国、中国、欧州などの開発が進むことで起こる問題がそうした帰結をもたらす可能性の方が圧倒的に高い。そのような状況下にあって、日本が反 AI 路線を探るよりも、AI 開発でキャッチアップを狙い続けつつ、国際協力などで重要な役割を果たし続け、リスク低減を狙う方が良い結果につながりやすいと思われる。

3. 8. 技術的的前提

いくつもの技術的な前提が存在している点も述べておくべき重要な論点だろう。たとえば投資の規模が AI の性能に大きく影響してきたことは事実だが、それが今後も続くのか、それともより小規模な投資で性能を高めるような競争が発生するのかについては明らかではないと思われる。利用可能なデータが文章や画像に関しては利用し尽くされる可能性があるため、その場合には投資規模を増大させることの効果が低減し、投資規模とは異なる次元での競争がより大きなウェイトを占めることになる可能性も考えられるだろう。(そうではなく、あるだけのデータからの学習をこれまで以上の投資規模で大量の計算資源を費やして非常に巨大なモデルを開発する方向に進展する可能性もあるのかも知れないが。)

文章やプログラミングなどのテキストを生成する AI の性能を評価するための数々のベンチマークで、優れたスコアを獲得してきているのは基本的には大規模なデータを使い、大規模な計算資源を使って学習した巨大なモデルだ。このような傾向が今後も持続するのであれば、巨大資本が AI 開発には有利ということになるだろう。だが、このような路線を進んでいくことが AI の発展経路として望ましいのか、また、それが社会のニーズに合致しているのかは自明ではないようと思われる。現在存在するこの種の AI は、人間用の数学オリンピックでトップクラスの正答率を誇り、司法試験に合格し、数十の言語間の翻訳をこなすことができる。多くの人間はこれらの内一つとして達成できないわけだが、このような AI を相手に旅行の計画を立てたり、悩み相談をしたり、アイディア出しの相手になってもらったりしているというのは、どうも強みと利用方法のミスマッチがあるようにも思われる。もっとも、教育界と同じで、AI を評価する際にも、正解のある問い合わせ投げかけて正答率をベースに評価することが採点者の主觀が入りにくいという意味では客観的な評価として通用しやすい。複雑で正解がない問い合わせについての AI の性能を考えると、そもそも人間による評価が困難で、AI が優れているのかそうでないのかもわからないことが出てくる。例えば複雑性は高いものの正解が存在するゲームの一種である将棋では、AI の指した手が人間の常識や理解に反していることがあり、人間による研究の後によい手であったことが判明すると

といったことも既に起きている⁶¹。

AIの学習方法が今後どのように進展していくかによっても、既存のデータから得られる「学習効果」は異なる。データ量の限界によってAIの大規模化に歯止めがかかり、別の次元での競争（例えばモデルの小型化、利用時の計算量の小規模化など）が進展するのか、投資規模が競争力を大きく左右する状況が続くのか、それがいつの時点まで続くのかは自明ではないように思われる⁶²。AI生成物を含んだデータからの学習はしばしばAI生成物を含まないデータからの学習に比べて結果が劣ったものになる（モデルの性能が低くなる）ことが知られているが、その度合いが軽減される可能性がないのかも、今後の研究にかかるつているところもあるようと思われる⁶³。

3. 9. 小括と考察

ここまで検討したような様々な論点を踏まえると、さまざまな前提条件が満たされるのであれば、日本の現行の著作権制度によって、生成AIの開発・普及を促進することは得策であるように思われる。だが、裏を返せば、そのような考え方を再考させるような状況が到来する可能性もまた、様々に考えられるということになる。

例えば生成AIの利用価格が何かの理由によって上昇し、広い普及が望めなくなったらどうか。その場合には経済的な恩恵もまた小さくなり、現行の制度を維持する理由はあまりなくなるかも知れない。（その場合には直接・間接に損害を受ける人というのも減り、制度を変更する理由もあまりなくなるかも知れないが。）技術状況が変わり、AI生成物を学習データに有効に使うことができるようになると、データを大量に使わなくても優れた学習ができるようになったなどの理由で制度を維持する意義が小さくなる日が来るかも知れない。

個別の論点についての検討から一步下がり、オープン化の概念との関係でこれらの論点をより抽象的に概観すると、他分野のオープン性の議論とはやや異なる議論になっている点があるように思われる。オープン性（参加性、学習データであれば様々な人が著作物の学習データとしての使い道を考案し、実際に利用してみることができる可能性）を高めることができが結局、より多く・より多様な主体の参加につながるわけではなく、特定少数の主体の影響力を増強し、その主体がオープン性の恩恵を受け取ることになるのではないか、という懸念があることが一つの大きな特徴と言えるようと思われる。

著作物のオープン化は、政策論議の文脈では典型的には「権利者団体」、映画スタジオや音楽レーベルなどの「メジャー」のような大きな資本が代弁する権利者の意向と、社会に薄く広く存在する多種多様な利用者の意向との対立構図の中で、前者の排他的権利（著作権）を弱め、後者により多くの影響力や決定権を与えることがオープン化であり、そこからは社会のメリットも、権利者のメリットも得られることがある、という点に注目する傾向があつたように思われる。AIをめぐる政策論議では「権利者」の側には権利者団体などに組織化されていないクリエイターを多く含んでいる。その中には創作活動から大きな経済的な利益を得ていない副業的なクリエイターなども含んでいる。もう一方の利用者の側には、巨大

61 古作（2023）、岡部、市川（2025）。

62 学習用データの枯渇問題については、専門家の間でも意見が一致しているわけではない。例えばMaslej et al. (2025)。

63 幡谷（2024）、Lee（2025）Shumailov, et al. (2024)。

な AI 開発企業も含まれている。そのような状況下でのオープン化は、特定少数に集中していた権限をより多様・多数の人に分散させ、そうした人が影響力を行使できるようとする、という構図では整理し切れない部分が多い。これが特定少数の強者による不特定多数の弱者の搾取になる、という整理は単純化し過ぎて不正確なものだが、オープン化が幅広い人の「エンパワーメント」になり、これまで影響力が弱かった人の影響力が強くなる、というような変化にはつながらないかも知れない。

もっとも、生成 AI の利用者を考えると、それまで文章や画像などの創作スキルを十分に持ち合わせていなかった人でも自分の発想や好みを手掛かりに、あるいは他の利用者の利用データを手掛かりに、さまざまな生成をすることができるようになる、という面がある。これは比較的単純なオープン化の効果に近い。検索や文書の読み解きが苦手な人にとってのサマリー機能や、外国語のスキルが低い人にとっての翻訳など、テキスト生成の AI を用いることで物事を知り、学ぶことのハードルが低くなり、幅広い人に手が届きやすい活動になる効果もあると思われる。

もうひとつ、生成 AI という形での著作物の利用は、社会にもたらすデメリットがかなり大きくなる可能性も議論されている。社会が広く被る損害と、著作物のクリエイターの被る損害と両方について非常に大きなリスクの議論がある。他の領域で議論されてきたオープン化で議論してきたデメリットと比較して突出して大きなデメリットが発生するかも知れない、という議論があるのは気がかりな特徴ではある。(あくまで発生する可能性の議論であって、発生するかどうかが見通せるわけではない点には留意は必要ではあるものの)。2010 年代中盤の AI による大規模失業論がその後の研究によって訂正されたように、こうしたリスクについての議論も、今後より小規模で具体的なリスクの議論に置き換わっていく可能性もある。だが、現段階でそうなることを確信するには、十分な判断材料はそろっていないように思われる。また、有効なリスク対策は日本国内での生成 AI の開発・普及を抑制することではないように思われる。

また、社会全体としてはメリットが大きいとしても、生成 AI の普及の恩恵を受ける人と損害を被る人がそれぞれ不均一に存在し、異なる職種・業種への転換や業務内容・業態の変更などが必要になるケースが出てくると思われるため、基本的には何らかの再分配政策が望ましいと考えられる。そこで次節では、学習データのガバナンスに再分配政策的な意味合いを持たせることの是非や、るべきデータガバナンスの制度設計について検討したい。

4. 政策とガバナンスによる対応

これまでの検討から得られた対策の方向性を総合すると、再分配政策は望ましいが、直接 AI 開発企業などからその原資を得るのではなく、自由で公正な競争に任せつつも、経済全體が得られる恩恵の一部を、特に損失を被る人々への支援に充てるべき、ということになる。

この方向性について更に肉付けを試みたい。

4. 1. 学習データの提供者への対価の支払いの現実性

学習データをめぐるガバナンスの一つのあり方として、学習データの提供者に対して対価を支払う方式が考えられる。現に存在している取り組みを参考に考えるなら、このような取り組みは、個々の画像を提供してくれる者への支払い、まとめた量のデータを提供する

プラットフォーム事業者への支払い、などの形をとる。このような制度のメリットは、学習データの無断利用に変えて、有償の許諾ベースの利用に切り替わることで、学習データとして無断利用される側の損失を補填できることにある。

このような制度の最大の問題は、AI開発に関する参入障壁を高くし、資本力の乏しい大学やスタートアップなどの事業を困難にする点にある。大企業は相対的に有利になる。

具体的な数字を少し参考してみると多少この困難のレベルを推しはかる助けになるだろう。画像生成系のAIについて見ると、例えば LAION-5B と呼ばれる画像データセットは画像生成系AIの学習データとして用いられることがあるが、規模としては 59 億点程度にのぼるとされる⁶⁴。DALL-E 2 は公開資料によると 6.5 億点程度を学習に使っている⁶⁵。DALL-E は 2.5 億点⁶⁶、Imagen は 8.8 億点、Stable Diffusion の初期バージョンは 23.2 億点程度を学習データとして利用している⁶⁷。

テキスト系データに目を向けると、Common Crawl や The Pile と呼ばれるデータが学習データとしてしばしば言及されるが、前者は月次のリリース単位で見るとおおよそ 30 億ページ前後を捕捉し、とりまとめたものになっている⁶⁸。Meta 社の Llama は当初、この Common Crawl の 2017 年から 2020 年までのものから取捨選択したものを主要データの一部として用い、分量的には全学習データの 3 分の 2 程度を占めていると報告している⁶⁹。The Pile は Common Crawl の一部を含め、質が高いと想定される多様なデータセットから成っている⁷⁰。規模はファイルサイズにして 900GB ほどで、発表当時の Common Crawl データセットのおおよそ 60 分の 1 程度と推定されている。それをウェブページに単純に換算すると、0.5 億ページ分程度ということになろう。

学習利用の適正価格は自明ではないが、仮に画像 1 点につき 10 円を 1 回だけ払うとして 25~250 億円程度の予算が必要となる。ウェブページに同様のレートを適用すると、5 億円 ~ 数百億円ということになるだろうか。画像一枚当たり 10 円、あるいはテキストページ 1 ページ当たり 10 円の支払い額は、おそらく自分の仕事が生成 AI によって脅かされる可能性を考えている人にとっては全く十分な額ではないと思われるが、AI開発企業はこれを払えるだろうか？ 2021 年の日本のベンチャー企業の調達額は総額（つまり全ベンチャー企業が調達した額の合計）で 8,000 億円、1 件当たりの調達額はレイトステージでも平均数億円規模であるとされる⁷¹。この規模の産業に、1 回の支払いであっても数十億円の負担を求めるることは難しいように思われる。日本の AI スタートアップの中でも大手と思われる SakanaAI は 2024 年に約 300 億円の資金調達をしたと報じられている⁷²。この程度の規模

⁶⁴ Schuhmann, et al. (2022).

⁶⁵ Ramesh, et al. (2022).

⁶⁶ Ramesh, et al. (2021).

⁶⁷ Imagen については Saharia et al (2022)に学習データについての説明がある。Stable Diffusion v1-1 は Rombach & Esser (2022)によると LAION-2B-en を用いているが、Schuhmann et al. (2022)によるとこのデータセットは 23.2 億点規模である。

⁶⁸ Common Crawl (undated)..

⁶⁹ Touvron et al. (2023).

⁷⁰ Gao et al. (2020).

⁷¹ 松井. 文谷 (2022).

⁷² 浦中 (2024).

の企業であれば可能性はあるかも知れない。

大手の AI 開発主体の資金力はどうだろうか。報道によれば、OpenAI 社は、2025 年 8 月に 1 兆 2000 億円超の資金を調達している。これは 2025 年を通じて 6 兆円程度を調達する計画の一部とされている⁷³。同社はモデル開発や提供だけを行っているわけではなく、データセンター構築に今後数百兆円規模の投資をする計画があることを述べているため⁷⁴、専業の AI 開発主体と比較するのは若干厳密さを欠くところもあるかも知れないが、投資規模には相当の違いがあるとは言えるだろう。同社の収益は 2 兆円近いと思われる⁷⁵。マイクロソフト社は 12 兆円程度のデータセンター投資を今年度に予定していることも報じられている⁷⁶。これらを含め、AI インフラへの投資額は大きなものになりがちであることが総務省の資料からも伺える⁷⁷。GPU の不足が頻発したことなども関係していると思われる。2023 年度の日本の主要通信事業者の投資規模や売り上げに対する研究開発費の割合を米中の大手テック系企業と比較すると、日本には例外的に年間 1 兆円程度の研究開発をしている通信事業者は存在するが米国では大手テック企業が 5-7 兆円程度を研究開発に費やしていることが伺われる⁷⁸。なお、日米以外で主要な生成 AI 開発主体としてフランスの Minstral 社を挙げることができるよう思うが、同社の資金調達額は 25 年 9 月のラウンドで 2 兆円規模であることが報じられている⁷⁹。画像系生成 AI の開発元として知られる Stability AI は 2024 年に 120 億円程度を、ドイツの Black Forest Labs は同年 40 億円程度を調達している⁸⁰。

このような資金力の差を考えると、一部の大手 AI 開発企業にはこうした支払いは可能かも知れないが、日本の AI ベンチャーなどは参入が非常に難しくなるのではないかと思われる。このような支払い負担が日本で導入されるとなると、おそらく日本の AI ベンチャーへの投資なども減るのではないだろうか。

実際に生成 AI の中で学習データ提供者に支払いをしている例として Adobe 社の Firefly が知られている。その金額については報道レベルでも確認できるデータはほとんど存在していないようだが社内文書によると 10 ドル以上の支払いを得た人は 100 万人強の適格者の中で 6.7 万人だったとの報道がある⁸¹。オンラインフォーラムの投稿では、2 回の支払い額の合計を投稿した画像数で除すると、1 枚当たり 3 セント弱と報告している投稿がある⁸²。日本には似た仕組みの事業として AC ワークス社の AC 写真 AI ラボというサービスが

⁷³ Sorkin et al. (2025).

⁷⁴ Heath (2025).

⁷⁵ Sigalos (2005)

⁷⁶ Newman (2025)

⁷⁷ 総務省(2025). (資料編「主要な AI 開発事業者の投資状況」

<https://www.soumu.go.jp/johotsusintoeki/whitepaper/ja/r07/html/datashu.html#f00028>)

⁷⁸ 総務省 (2025) (図表 II-1-1-16 日本大手通信事業者と GAFAM・BAT との研究開発費の比較)

⁷⁹ Clark (2025).

⁸⁰ Sriram & Hu (2024), Nuñez (2024)

⁸¹ Kim, E. (2023).

⁸² jar (2024).

存在している⁸³。画像生成AIの学習データ画像を提供した者に対し、生成画像1枚につき3点までの影響力の大きな画像について対価を支払うもので、支払額は1-3円程度である。この生成AIの学習データは1,000万枚程度と小さいが、どの程度の利用量があり、それが学習データ画像1枚あたり平均いくらの支払いになるかなど具体的なところは開示されていないと思われる。

4. 2. 支払い対象者の限定とそれが生む追加的コスト

もっとも、仮に支払をする場合でも、外国人はそもそも手続きをするのにお金がかかりすぎるため、10円や100円程度のお金であれば受け取ろうともしないであろう。学習データ用の画像はスナップ写真や製品の説明画像など、様々なものを含んでいるだろう。これらの中から、たとえば写真やイラストなどとして販売されている作品だけに限ることにし、いわばダメージを被る可能性がある人に対して補償を行うような制度にすることも考えられる。このようにして、仮に支払額が500分の1以下になり、1000万円程度になったら支払えるだろうか。日本のAI開発企業の中にもこの程度の額であれば支払える企業は多くあるかも知れない。大学などの研究機関にとっては厳しい額になっているかもしれない。同様に、文書系の生成AIについては、脚本家やエッセイスト、ライター、ジャーナリストなどに限定して対象にすることが考えられるのではないだろうか。ただし、1000万円という規模から容易に想像できるように、これは生業に大きなダメージを受ける人を広く補償するための金額としては全く不十分なものになるだろう。

金額を縮小する場合には、もう一つ別の困難も発生する。趣味として（ただの消費活動の形で）写真やイラストを公開している人々や小遣い稼ぎ程度の趣味と実益を兼ねるような形で公開している人々の分について補償がされることになるとして。この辺りの対象者のくくり方や、実際にどのような形で適切な対象者だけを特定し、支払いを行うのか、ということを考えると、実際に支払う額以外にもそもそも支払い手続きに多大な費用がかかることが予想される。

その点に対処するべく、支払いをめぐるコストを圧縮し、中小のAI開発者でも支払えるようにするために、関連の業界が協力して補償金の収集・分配の制度を立ち上げ、運営することも考えられる。その設計なども含めた手続きや、その前段の合意形成等には数年程度の時間がかかると思われる。その間に生成AIの技術動向が変動する可能性なども考えられるところで、そこからもガバナンスのあり方としての難点を感じる。

もう少し実効性が高いのではないかと思われるのは、投稿型サイトの類、大手のプラットフォームがこの仲介を行う制度だ。このようなサイトは、投稿者とアカウントを通じてつながっており、場合によってはサイトの利用料を徴収するなどの金銭的取引をしている場合もある。このような大手サイトがAI開発企業から金銭的支払いを受ける代わりにデータの提供を行っているというニュースは散発的ながら存在している。同じコンテンツを複数のサイトに投稿している者は複数回の支払いを受けることになるなどの明らかな問題があり、また、中小規模のプラットフォームはAI開発企業から興味を持たれにくいために、大手投稿サイトが投稿者の獲得競争上有利になるのではないか、という問題もある。だが、中小サ

⁸³ IT Media (2025b).

イトが連合体を作ることが許されれば、その点は克服できるかも知れない。

この方式の問題点は、大手開発者が有利になりやすいと思われる前出の点である。例えば Reddit は Google や OpenAI との契約を通じてデータの提供をしているとされるが、ロイターの報道では Google の支払い額は 1 年間に 6000 万ドル程度であるという⁸⁴。当時の円ドル為替相場で換算するとおよそ 90 億円程度になる。金額が公表されないことが多いが、Shutterstock 社が 2500 万ドルから 5000 万ドル程度のデータ提供契約を複数社と締結している、Freepik 社は画像 1 点当たり 2 セント～4 セント程度（数円）で自社保有の 2 億点程度の画像を提供している、といった報道もある。（ということは 5～10 億円程度の売値になっていると思われる⁸⁵。）このような報道を割り引いて受け止めることにして、例えば、ある程度高い価格での提供であるからこそ公表され、報道されている可能性がある、実際の市場相場は報道されていないもっと安価な取引を含むはずだ、と考えることにしても、やはり参入障壁としては無視できるレベルではないと思われる。

4. 3. 支援対象の範囲等に関する疑問

関連する論点として、生成 AI の普及と共に損害を受けることになるであろう人々が政策的な救済・支援策の対象として特別なのか、という疑問がある。

例えば、10 年近く前に遡って、生成 AI とは違うが似たようなケースを考えてみることができる。生成 AI が注目を集めようになる以前には、現在では識別系 AI と呼ばれるような機械学習（中でも深層学習）の応用可能性について多くの議論が行われ、そこでも幅広い応用可能性、経済的な恩恵や損害が議論された。この領域でも、AI にデータを提供することで職を失うリスクが高まる人がいた点について、どう考えるべきだろうか。例えば、放射線科医である。機械学習の分野で世界的権威であるジェフリー・ヒントン氏は放射線科医が AI によって代替されることを予言した。2016 年のことだった。予言では 5 年か 10 年以内に深層学習が放射線科医を代替することが「完全に明白」とのことだった⁸⁶。このような AI も学習用データを必要とし、そのデータが単に検査画像だけでなく放射線科医の判断した内容などを含んでいることが重要となる⁸⁷。ではこのような立場にある放射線科医はデータの提供を拒む、相当の許諾料を要求する、AI 開発の遅延のために働きかける、といった行動をとるべきだろうか。また、著作権法上はこうした医師の判断内容やその記載などは著作物としては保護されないことがほとんどだと思われるが、著作権を拡張して、こうしたデータも保護の対象とし、AI 開発企業に対して許諾料を請求する法的根拠を与えるべきだっただろうか。難しい問い合わせだが、昨今の生成 AI とその学習データ利用から直接的に損害を受けるかも知れない人だけを特別に著作権法で保護するべきなのか、という点はこのような「ほかの AI」を考えると疑問が大きくなる点である。もちろん、これを更に拡張して「AI は特別なのか」と問うこともできる部分がある。ウィキペディアが台頭したことで、既存の百科事典や用語事典などの中には事業継続が困難になったものがあると思われる。（マイクロソフト社の所有下にあったエンカルタ百科事典や、その元になったと思われる Funk &

⁸⁴ Tong, et al. (2024).

⁸⁵ Paul & Tong (2024).

⁸⁶ Hinton (2016), Byju (2024).

⁸⁷ 町田 (2023).

Wagnalls百科事典を考えることもできるだろう。ブリタニカなどの百科事典の執筆・編纂者にとっての収入機会の逸失を考えることもできるだろう)。ブログやオンラインメディアなどの台頭によって衰退に拍車がかかったと思われる雑誌業界で廃刊になった雑誌を想定することも出来るかも知れない。上述のマイクロストックフォトの台頭と共に、ストック写真からの収益が大きく減ってしまったであろうプロの写真家を想定することもできるだろう。こうした人々への補償は行わないものの、生成AIの学習データとして利用されたクリエイターに対しては補償を行う、ということをどう考えたらよいかは自明ではない。

ちなみに、ヒントン氏の予言は2016年の時点で、その後もそうした説は珍しいものではなかったようだが、結果として今日まで著作権法は拡張されず、放射線医はAIを利用はしている例も少なくないものの代替されてはいない⁸⁸。加えて、米国ではAIへの注目度が上がった時期には、放射線科医のなり手が減った⁸⁹。おそらくはそれも手伝って放射線医不足が深刻で、燃え尽きも非常に多いためAIを必要とする状況になっているとも言われる⁹⁰。日本でも検査機器の高度化に伴いデータ量が増大しており、担当医の負担が増加しているとの指摘もある⁹¹。こうした経緯を見ると、ガバナンスや政策の失敗リスクも気になるところだ。つまり、代替に関する説が外れる可能性があり、AIの開発が進んでも性能面で想定以下の伸びしか出ない、普及面では現場の業務フローとの親和性、その他の理由で想定以下にとどまる、などの理由でそもそも事態が大きく変動しない可能性がある。また、AIが脅威になるのか必要になるのかについても不確実性がある場合もある。こういう状況下では、「直接的に被害を受けるであろう人を事前に保護する」ようなガバナンスや政策が失敗すること、政府による政策であればいわゆる「政府の失敗」につながることは、驚くには値しないように思われる。生成AIをめぐっても、こうした状況はあるだろう。

生成AIの領域はかなり技術の先行きが見通しにくい領域のように思われる。そもそも大規模言語モデルの性能がここまで高いものになると自体が専門の研究者にとっても想定外の出来事であったとしばしば言われ、半年先の展開もまた非常に予測が難しいとする研究者もいる。

こうしたことからも、生成AIの活用によってさまざまな職場で効率化が起こるのであれば、それを止めることはせずに、転職や業務内容変更などが必要になった人を支援するよう、「損害」が発生した際の支援・救済制度を充実させる方がよいのではないか。その場合には、データの提供者であるかどうかではなく、また、生成AIであるか(生成AIとは異なる)識別系のAIであるか、その他の理由によるかなどを区別し、ある技術の被害者は支援するが別の技術の支援者は支援しない、という風に扱いを変える必要もないのではないかと思われる。

出産・育児であったりうつ病など健康上の理由であったり、家族の引っ越しであったり、職場の人間関係であったり、待遇面への不満であったり、キャリアパスに関する考え方であったり、会社の倒産や人員整理であったり、契約の終了であったり、人は様々な理由で離職し、

⁸⁸ 日本貿易振興機構 市場開拓・展示事業部 海外市場開拓課 ニューヨーク事務所 (2022).

⁸⁹ 横山, 桶川 (2021).

⁹⁰ 中田 (2024).

⁹¹ 桶川 (2024), 町田 (2023).

失業者になるが⁹²、こと AI との関係で議論されることが多いリスクリング政策については、AI 普及との因果関係が認められる失業・離職などに対象を限定する必要はない面が多いのではないかと思われる。因果関係の確認自体が非常に大変になると思われることが主な理由だ。どういった人材が離職するのかを把握することはリスクリングに関する制度設計や運用には有益だが、それも AI 起因の離職にのみ限定する理由はないようと思われる。

4. 4. 著作権制度の扱い

著作権については、大きく 2 つの主要な論点があるよう思われる。ひとつは生成 AI のような技術の普及を受けて、作風などこれまで著作権法で保護してこなかった領域に保護を広げるかどうか、という点である。もうひとつは学習データ利用を可能にしている著作権法の規定（権利制限規定）の可否や規定ぶりである。

1 つ目の論点については、作品に人々が求めるものと、著作権法が保護する対象のズレだ。著作権法は創作的表現を保護するもので、作風と呼ばれるものにはあまり及ばない。法律の世界の言葉なので「創作的」も「表現」も一般人が考えるような意味とは少し違っていて、やや広いが、それでも作風と呼ばれる領域にまではあまり及ばない。

では人々は個々の作品の特定の創作的な表現をどこまで欲しているのだろうか。その特定の作品でなくとも、似たような作風の別の作品でもよかつたのではないか。一部の著名な作家については、「この作家が創作した本物の作品である」という類の性質を持つことが重要で、似ているだけの作品には価値がないということになるだろう。だが、数多く存在する作家のほとんどについて、あまり細かく把握も区別も出来ていない人がほとんどなのではないだろうか。

作品の表現部分は保護され、作風部分は保護されず、両者はこれまで分離できない形で流通してきたので問題がなかったが、生成 AI はこの分離をある程度可能にしてしまったために、かなりの人々が、著作権で保護されていない部分で満足してしまうことになるのではないか。

では人々が価値を感じる作風にも著作権の保護が及ぶようにするとよいのではないか、という考え方もあるところだが、作風の保護をしてしまうと、かなりの程度文化が停滞してしまう可能性も考えられるのが悩ましいところだ。絵画のような視覚芸術でも、音楽や文学でも、流派や主義のようなものがあり、ある作家のある時期の作風が別の作家に影響を与えるということはしばしば起こる。これを著作権で保護してしまうというのは、かなり広い範囲の模倣を禁止してしまうことになり、いわば流派の形成も（許諾を伴わなければ）できない、といったことになりかねない。

そして、著作権の領域では、このような、創作的な表現ではない要素について保護を求める要求は非常に多く存在している。生成 AI とも関係が深いところで言えば、ジャーナリズムの領域では、事実の情報についてそれをスクープした報道機関に独占権を与えてはどう

⁹² 日本における離職理由は総務省(2023)の p.26 表 8-2 や厚生労働省(2025)の p.23 付表 3-1 などから伺える。AI などの技術が直接の理由になっている選択肢はそもそも存在していないが、技術の普及と共に変化する経営環境に会社や自分がついていけないことが離職につながる、といった事情がある場合でもこうした統計の複数の項目に反映されることになると思われる。

うか、という議論がある。確かに報道機関は専門家を動員して、時間やお金をかけて調査を行い、独自の重要な事実を発見するようなことがある。事実自体は保護の対象ではないから、スクープの内容は、ほかの報道機関によっても報道されることになる。それではスクープをした報道機関は十分な投資回収ができない、そうするとお金のかかるような調査報道への関与や、本当に重要な発見につながるかどうかが不明瞭な件について調査をしていくようなリスクは減らざるを得ない。これはこれで社会の損失となる。ところが事実の情報についての報道の権利を独占させると、ニュースの流通が停滞するから、こちらもまた非常に大きな損失を招くことになるだろう。

ほかにも著作物であるという主張が法廷で退けられて、著作権法の保護対象ではなくなったものの、努力や投資の結果生み出されたものはいろいろある。学術的な研究の結果到達した概念の定義、スポーツのルールなどもそこに含まれる。こうしたあれこれが「著作物」になるとすると著作権を持っている人から許諾をもらわなければ複製や改変ができなくなる。他にも実用品（椅子の形状、公園の滑り台の形状など）については長年著作物ではないとされる傾向が強かったものが認められる例が出てきて、その基準がまだ明確に統一されていない状況にあるように思われる。こうした文脈を意識すると、生成AIに関連して著作権の保護範囲を拡大することが適當とは言いにくいと思われる。

事実情報は文章生成をするような生成AIでは頻繁に利用され、それ故にウェブサイトへの訪問者が減少するということは起こっているように思われる。報道機関の弱体化は民主主義にとっても大きなリスクになるとは思うが、他方で事実情報に著作権による保護をかけるというのは社会全体にとっての被害が甚大になり過ぎるだろう。絵画やイラスト、写真などの作風については、それに比べると被害が少ないかも知れないが、流派形成やほかの作家からの影響を受けた創作などが厳しく制限されるほど文化の停滞につながる面はあるだろうと思われる。

更に言えば、受け手が価値をおいているのは、作風ですらない場合も多くあると思われる。つまり、ある作家Aのある作品Xを楽しみたい、ではなく、また、ある作家Aのような作風の作品を楽しみたい、でもなく、例えばアニメ風とか、リアリズムとか、昭和レトロとか、優しい感じとか、スタイリッシュな感じとか、非常に抽象的なレベルでは好みがあるかも知れないものの、より具体的にはそれほど細かいこだわりがない場合も多いのではないか。これは画像生成AIサイトで人々がどのような画像を生成しているか、どのようなプロンプトや追加学習モジュールを使っているかなどを観察すると気が付くことだが、細かいこだわりを持っていない人はかなり多いという印象を持つ。（反面、特定の作品の特定のキャラクターに強く興味を持って画像生成をしている人なども多く存在しているという印象も持つ。）そうだとすると、仮に画風と呼ばれる領域をある程度著作権法で保護することができたとしても、人々が画像生成AIの生成物を使うようになって写真やイラストに対する需要が減少する、という事態は大きく変わることはないかも知れない。

もっとも、著作権制度についてはもうひとつ主要な論点があると考えられる。上に2つの論点として言及したものであり、本稿の中心的な関心でもある。生成AIと著作権の関係では、そもそも似た画風の生成を禁止するかどうかといった「生成」以前の工程に問題がある、という見方ができると思われる。というのも、アウトプットの生成に関しては原則侵害がないと言えると考える場合であっても、学習用データを収集・加工する段階ではそのデ

ータの複製をしているためだ。このようなデータの収集・加工部分だけを原則違法にしておけば、当面の生成 AI の研究開発は停滞し、許諾を得られるデータのみから学習することになる。例えばオプトアウトの可能性を確保することでも、少なくとも当面は、学習データの利用しやすさを大きく低減させることができる。EU では、識別系の AI を想定して学習データの利用をしやすくしたもの、生成 AI にもこれを適用することが妥当なのか、かなり議論があることも伺える⁹³。

別の角度からは次のようにも言えるように思われる。一方では、著作権制度は既存のビジネスモデルを保護することを目的とするものではなく、技術環境の変化によって既存のビジネスが成立しにくくなる時にそれを救済する形で既存の文化産業を特に優遇する必要はない。また、その技術環境の変化が著作物の創作的な表現ではない部分の利用（事実情報や作風など）に留まるのであれば、著作権法の基本的な原則を変更してまでそれを禁止する必要はない。他方で、この変化は 2018 年の著作権法の改正を背景に可能になっていると見ることもできる⁹⁴。しかも、大手の開発者以外は、おそらくは十分なライセンス料を支払う能力を持ち合わせていないと思われる（これについては後述する）。著作権法改正がなく、無断での学習データ利用が著作権侵害になっていたのであれば、多くの AI 開発企業は日本のデータを利用できず、無断で利用した企業が訴訟を通じて破産することも考えられただろう。無断利用を禁じるという著作権法の基本的な原則を変更してまでそのような十分な支払い能力を持ち合わせていない新興技術を支援する必要はないのではないか。ましてや、それによって既存の文化が大きく衰退してしまう可能性もあるとなれば、慎重に検討すべきではないか。そのように考えるなら、せめてオプトアウトを可能にする、あるいは補償金制度を創設するといった対応が適当なのではないか、とも考えられる。

もっとも、これについても反論が考えられるところで、「無断利用を禁じるという基本的な原則を変更」したのかというと、学習データ利用をめぐる著作権法改正は、著作物の「享受」をしない利用について無断での利用が可能だという規定を導入しているので、むしろ著作権法の原則に含まれている内容を明文化しただけに過ぎないと見ることも出来るようと思われる⁹⁵。

そもそも現状の制度は、柔軟な著作権制度を設けることで、イノベーションを推進することが望ましいとの観点から設けられたもので、与党による方向性の明示や知財戦略本部の決定なども含む議論の過程を経ている⁹⁶。AI に限らず、個別の規定を導入することで、次々に登場する様々なサービスや著作物利用行為を合法化するよりも、著作権法が本来保護するべき利益である著作物の享受利用に関わる部分以外の著作物利用について、柔軟に適法性を判断できる制度を設けることが日本のために望ましいという判断が背景にある。AI を後押しするのもそのような取り組みの一環だった。その経緯を考えると、既存の著作物を

⁹³ Lucchi (2025).

⁹⁴ 文化庁 (2018a), 文化庁著作権課 (2019).

⁹⁵ この考え方について前田 (2024) がある。「AI の促進のために著作権の保護の後退を強いているわけではない」とも述べている。

⁹⁶ 与党の関与については自由民主党政務調査会 (2016)、当時の文脈を知る手がかりとして大野 (2016)、榎原 (2016)。行政側の経緯については、文化庁 (2018a, pp.1-11; 2018b, pp.1-8)。

活用することで新しい技術開発や産業が生まれることは制度で想定され、望まれていたことの範囲ではあるだろう。それが著作者の利益として保護するべき部分（著作物に表現されている思想や感情を享受すること）を保護し続けている点もその通りである。その文脈では、立法の過程で想定していたさまざまなイノベーションの一つが、この柔軟性をベースに成立していると言える。それが日本にメリットをもたらすようなイノベーションになっている、という点でもおおむねその通りと言えるだろう。

とはいっても、このような形で生成AIが急速に台頭することが予期されていたわけではないだろう。生成AIは2022年のChatGPTの爆発的な普及以前にも存在しており、いわゆる識別系のAIを中心とした政策論議の中でも取り上げられる機会はあったため、完全に想定外だったということはない。また、生成AIに対して、政府や与党は基本的にこの開発や利用を後押しすることを重視した政策を打ち出しているが、2018年の法改正に至るまでの議論の過程で、2022以降には今日のような生成AIの高速な高度化や普及といった事態が起こることを意識していたら（その是非はともかくとして）別の法制度になっていたのではないか、という疑問は残る。

米国では、同じ柔軟性を確保したフェアユース規定により合法であると考えられるものの、これは法廷の判断を待つ部分も多く、本稿執筆時点ではそれほど判例動向が明確になっていない。違法にコピーされている著作物（いわゆる海賊版など）を収集してそれを学習データとすることについては違法とする意見が出ているなど注目に値する個別的な論点が出ているものの、例えば最高裁判例が出るなどしてどのような範囲の学習データ利用がフェアユースに該当するかがおおむね明らかになった、という状況になるにはまだ時間がかかるように思われる。

EUでは基本的には権利者がオプトアウトできる余地を確保しており、AI法および著作権指令を通じて生成AIの学習データ利用を拒否できるように規定が盛り込まれている⁹⁷。ただし、これについては国際的に見るとAIの開発を不利にする制度であるという指摘もある。例えば、日本のような制度がAI開発にとって望ましいが、著作者が新たな創作活動を行うインセンティブを考えるなら、オプトアウトは禁止するものの補償金などを支払う仕組みを創設することが望ましいとする論などもある⁹⁸。

現時点では判断が難しい部分があるが、日本が米国やEUに合わせるためにメリットが

⁹⁷ デジタル単一市場における著作権指令 第4条(3)、およびAI法第53条(1)(c)が汎用AIの学習データのオプトアウトを可能にする義務を定めている。AI法の53(4), 55(2), 56条の規定を受け、European Commission (2025)の汎用AI向け実践規範(The General-Purpose AI Code of Practice)では、AI法に定められた義務の履行法の一例を明確化している。ここではオプトアウトについてIETFのRFC 9309に言及しつつ、robots.txtによるオプトアウトに対応することが義務の履行の一部をなすとしている。なお、この実践規範の性質上、この規範に従わなければ即EU AI法に違反することになるわけではなく、この規範に従っていれば規制当局は法を遵守していると前提するが、それ以外にも遵守する方法がありうるということになっている（いわゆるセーフハーバー的なアプローチになっている）。従って、EUはIETF RFC 9309の定める robots.txtへの対応を義務付けられたとまでは言えないが、それに近い状況が存在することになる。また、Google, OpenAI, Microsoft, Mistral, IBMなどAI開発企業がこの規範を採用する旨を表明している。

⁹⁸ Peukert (2025).

あるかどうかを考えると、補償金の支払いや許諾料の支払いなどといった金銭的な支払いの仕組みを意味のある額と共に確立することが難しいという上述の前提に立つなら、メリットは大きくはないように思われる。それよりも学習データの利用を容易にしつつ、性能のよい生成 AI の迅速な開発やその普及を後押しし、そこから社会全体が得られるメリットを前提に、リスクリミングのような制度を充実させる方がよいように思われる。ただし、オプトアウトのような拒絶を可能にすると、どれだけのクリエイターが創作の意欲を維持できることになるのかは現段階では判断材料が乏しく、文化全体への大きな影響・損失があるかどうかについても不明瞭な部分が残るように思われる。

4. 5. 自主的な制度形成とガバナンス

本稿のような検討の文脈でもう一つ注目に値すると思われる提案がある。現段階では草稿的なものに過ぎないが、米国に本拠をおくクリエイティブ・コモンズが提案している CC Signals（以下 CC シグナルと表記する）というものだ⁹⁹。今後の展開については不明瞭であることから、詳細部分について立ち入って議論することは本稿にとってあまりメリットはない前提で概略的な特徴に着目すると、法的強制力を持たせず、民間の自主的な秩序形成を試みている点と、ある種の互酬性や公正性を志向している点、その互酬性をいくつかの種類に限定していることについては、興味深く、また、独自のメリットもあると思われるため、そうした特徴を含め以下にもう少し説明し、メリットを述べてみたい。

この制度は、まとめたデータセットを保有する者（上述の Reddit やいくつかの画像系サイトはこれに当たると言えるだろう）が、データ利用を望む者に対して、どのようなリターンを求めるかの意向を表明できるものとしてデザインされている。

クリエイティブ・コモンズといえば、クリエイティブ・コモンズ・ライセンスという著作権に関わるライセンスがあり、こちらは法的に有効なものとなるようにデザインされているが、CC シグナルは法的強制力を持つわけではない意向の表明となっている。意向の内容は有限の要素に限定されており、データを提供したコミュニティへの還元や開発された生成 AI をオープンソースとして公開することなど、データ提供者個人などへの直接的支払いとは違う形での利益還元が含まれている点にも特徴がある。

この背景には、クリエイティブ・コモンズが重視するコモンズの維持可能性や公正性への関心があると説明されている。本稿の文脈で言えば、生成 AI の開発主体がクリエイター（学習データの「提供者」）から恩恵を受けながらもそれを還元しないという構図があり、それが続くようなら「コモンズ」が維持不可能になってしまう、というのだ。例えばそれは新しい創作活動が減ってしまうこと、創作されたものがネット上で公開されなくなってしまうこと、更にはこうした作品がクリエイティブ・コモンズ・ライセンスのように広く公衆に利用を許可するライセンスの下に提供されなくなってしまうこと、などを含んでいると思われる。

このような制度は、一方では生成 AI がもたらすであろう恩恵を減らさないようにするという点では優れており、具体的な支払い額などではなく、強制力もないことから、中小規模の AI 開発者にとっても、（負担可能な額をはるかに超える料金などではなく）身の丈にあ

⁹⁹ Creative Commons (undated)

った利益還元の可能性を与える。うまく機能すれば、少なくとも一部のクリエイターにとっては一方的な搾取があると感じて創作のインセンティブを削がれるという状況を解消するもののように思われる。

また、このような意向に応えて開発される生成AIと、これを無視して開発される生成AI、あるいは完全に無視しているわけではないが、完全に意向を満たしているわけでもないような中間的な生成AIなどの間の競争が起きやすいこと、また、データセットの提供者にどのような意向が表明されるかを概観することでどのような規範が多くの支持を得るかについてこれまでとは異なる情報が手に入ることもメリットがあるようと思われる。より強制力を持つ制度を設計する場合でも、どのような制度が最適かを判断する材料があることが望ましく、そのような判断材料の一部をこののような制度が提供することになると思われるためだ。

互酬性の種類をほんの数種に限定していることは、クリエイティブ・コモンズ・ライセンスでおなじみの手法で、おなじみのメリットがあるようと思われる。データセット毎にあまりにも利用条件などが異なっていると、これを満たすための労力が膨大なものになりやすい(ライセンスの多様性は、組み合わせ利用の際などに問題になりやすい)。これに対して、あらかじめ限定された少数の種類の互酬的な行動が示されるのであれば、それに応じてとするべき行動も比較的単純に検討できることになる。示される意向の種類が何であるかは機械判読可能(マシンリーダブル)なものとなるようにデザインされている点も、同様の観点からメリットと言えるだろう。

民間の自主的なガバナンスであることにもメリットがあるようと思われる。法改正は簡単なものであっても2-3年を要することが多いが、民間の自主的ガバナンスの仕組みであればもう少し状況に即応したルール作りなどができる場合もあるだろう。これはメリットと言えるようと思われる。もっとも、明らかな短所や限界もある。民主主義的な手続きを経ていないことで正統性が欠けやすい、従わないプレイヤーに対する罰を与えにくい、などは明らかな短所であろう。ただし、互酬的な行動が広く求められるのであれば、罰を与えるためのアクションをとらずとも従わないプレイヤーの人気が落ちる、批判が多く出て評判が落ちる、といった不利益が発生するかも知れない。

生成AIに関しては国際的な議論がされることにはメリットがあるだろう。国際的な制度形成は、国内法の改正よりも更に多くの時間がかかることが多くあるようと思われるためだ。ただし、国際的な議論については言語の壁があるため英語話者の少ない国やグループの意見が届きにくいなどの欠点もあるだろう。また、学習データ提供者に対して金銭を支払う仕組みを一国内で、民間主導で立ち上げる場合と比べると利害関係者が多いために利害関係の調整に時間がかかることになるリスクもあると思われる。また、国家でなく民間が主導する場合には、国家や国際機関に比べると人的リソースが限られ、長期的な持続可能性の担保が難しくなる場合もあると思われる。

個別のクリエイターではなく、データセットを保有している者がこれを採用するという想定は、そのようにしておいた方が生成AIの開発者にとっても無視しづらく、従うインセンティブの高いものになる、という点と、この仕組みが robots.txt に組み込まれることを想定してデザインされていることとも合致する。この robots.txt ウェブ上のデータを収集するために生成AI開発者などが使うクローラーに対して、ウェブサイトが収集の可否など

を伝えるのに使われるもので、検索エンジンに対する意向表明に活用されてきた。日本では著作権法によって、これに従うことを条件に、検索エンジンがウェブサイトのデータを収集することなどが著作権侵害にならない、という制度ができている。

このようにメリットも、限界も考えられる制度だが、短期的には生成 AI の開発者、データ提供者、生成 AI の利用者のそれぞれから賛同者が十分に集まるかどうかが課題となるだろうと思われる。この CC シグナルとは関係なく、現状では、生成 AI の開発企業は、robots.txt に示された意向に従っている場合も、そうでない場合もあることが報道からは伺える。この背景には開発競争の圧力があるのではないかとも思われるが、そうだとすると、賛同する開発者の生成 AI がより人気を得るようになるか、というのが次の課題となるだろう。例えば現在では生成 AI を使ったコンテンツは、ネット上で批判の対象になることがあるところ、このような自主規制の枠組みにのっとった生成 AI を利用していれば、批判が弱くなり、支持者も増える、というような違いがあるだろうか。逆にそのような利用者やネット言論を通じた影響力が強すぎれば、結局はそれに応えられる者だけが勝ち残れるという環境であり、データに料金を払わなければ開発者として機能できない、といった場合と大きく違わない結果になる可能性もある。

このように成否や効果について不確実な点は多いものの、うまくいけば AI 開発を遅延する効果は少なく、学習データとして使われる作品を生み出すクリエイターの内ある程度意味のある一定程度の人々のインセンティブは大きく削がない、という形になると思われる。

4. 6. 小括と考察

以上のような検討から、生成 AI の開発の際に学習データとして著作物を利用され、普及と共に収益機会の減少やキャリアの先行きの不透明化、失業などの損害を被るかも知れない人々に対して、AI 開発者から意味のある額の補償金を支払える可能性は少ないとされる。また、その範囲の設定や設定した範囲の人だけに適切に支払いをする仕組みの確立なども相当困難があると思われる。著作権法を改めて、作風に当たる部分も保護することにも大きな困難がある。EU や米国は日本に比べると学習データの扱いに関する著作権法上の位置づけが確定していない（米国）か、AI の開発を高額・困難にしやすい（EU）制度になっているが、それらに合わせるよりも、生成 AI のメリットを大きくし、それをいわば「原資」としつつリスクリングなどの支援策を充実させることの方がどちらかと言えば優れた政策オプションであるように思われる。文化への影響やクリエイターの創作意欲への影響、生成 AI の技術的な展開など多くの要素によるところも多いため、今後も調査を行うことや状況の推移に注視することも重要になるだろう。

こうした制度設計に関する主要な選択肢についての検討から一步下がって、より抽象的に概観すると、最後に触れた自主的な制度形成とガバナンスの取り組みをクリエイティブ・コモンズが主導しようとした点は、生成 AI がもたらしている状況を理解する上では示唆的であるように思われる。というのも、クリエイティブ・コモンズは、本稿で述べたような「狭義のオープン性」を推奨することが多く、そうでなくても広義のオープン性を推進し、クリエイティブ・コモンズ・ライセンスもその両方の側面を持っていた。その中には一部、一方的な搾取などを牽制するようなライセンスも存在していたが、それが強く推奨されてきたわけではないように思われる。（たとえばフリーソフトウェア財団は、GNU Public License

(GPL) がそうであるように、ある種の互酬性の担保を非常に重視していることと対照的である。GPL では、ソフトウェアを改変することは許諾されているが、改変してきたソフトウェアもまた、他人に対して GPL で提供することが義務付けられているため、互酬性の強制が起こりやすい)。

既に第 3 節の小括と考察で述べた点だが、オープン化はインターネットやそれを含む情報処理・通信技術の発展や普及と共に個人など影響力の弱い主体がエンパワーされ、従来はできなかつたような活躍ができるようになることを後押しする一つのムーブメントであった。参加の機会が増えることでその個人が持っていた知識やアイディアや意見がよりよく、「民主的」な社会を作ることにつながるという構図が成立しやすいためだ。

クリエイティブ・コモンズの黎明期は、インターネット上でファイル共有が盛んであり、著作権制度不要論なども存在していた。レコード業界や映画業界の業界団体がファイル共有ソフトや、動画共有サイトなどで起こる違法共有への対策として様々な訴訟も起こしていた。クリエイティブ・コモンズは、ネットを通じた共有がより参加的な文化を作る可能性にたびたび着目しており、人々が共有し、自由に使える作品を増やすことで、そのような参加型文化が豊かになるという想定の下に進められてきた。著作権の強化などにもしばしば反対しており、自由な利用ができる著作権制度を評価する傾向もあった。これは、著作権によって保護される主なプレイヤーが巨大なコンテンツ産業である状況では、そのような巨大なプレイヤーに対して、個々のユーザーの自由を擁護し、その自由の中から、参加型の文化が生まれてくることを後押しするものになっていたように思われる。

生成 AI をめぐる状況は、これとは大きく異なる側面も持っている。巨大な産業は著作権者側だけでなく、著作物の利用者側により、これに反対する側には、著名ではない人も、プロではない人も含めた個人クリエイターがいる。この部分では、著作物の自由な利用（すなわちオープン化）の推進と、個人など影響力の弱い主体のエンパワーメントが対立する構図になっているように見えることがある。

ちなみに、オープン化の推進と、個人などのエンパワーメントが一致しない、むしろ相反するのではないか、という構図は生成 AI に固有のものではない。プライバシーや個人情報を関連して似たような議論がある。というのも、自分の私生活に関わる情報を公開し、自由に誰でも使えるようにすることは、一面では多くの人から有益な情報や助力や機会などをもらえることにつながり、「自分がコントロールできる度合いは低くなるが、結果としてはより多くの協力者が得られ、よりよい決定ができるようになる」といったオープン化の典型的な恩恵を受けることにつながる可能性もあるが、反面、巨大な企業や国家に多くの個人情報を把握され、自分のアテンションや購買意思などを操作され、搾取されることにつながるのではないか、という可能性もあるためだ。この後者のリスクは、近年では巨大 IT 企業への警戒という形で広く懸念されるようになって来ている。

クリエイティブ・コモンズは、黎明期に著作権法廃止論などの急進的な立場は採用せずに、既存の著作権に根差しつつも、より個人をエンパワーできる方向を志向してクリエイティブ・コモンズ・ライセンスを開発・振興したと言えるが、生成 AI についても、これを根本的に否定するような急進的な立場は採用せずに、著作権法の大きな転換も提唱せずに、より参加性の高い文化が生まれる余地を確保する方向を志向して CC シグナルを提唱していると言えるように見える。2000 年代初頭の発足当初には、人々が利用したいと考える著作物

は大企業や著名なアーティストによって生み出されていた度合いが高いが、四半世紀ほどが経った今日では、インターネット上に大量に存在する様々なクリエイターの様々な作品が価値を持っており、クリエイターエコノミーと呼ばれるような各種市場も存在しており、従って、著作物のより自由な利用を促進することだけが個人などのエンパワーメントである度合いが低下していると思われる。また、そのような小規模クリエイターの活動を支えるプラットフォームの中には巨大な企業が多く含まれ、より自由な利用を促進することが個々人のエンパワーメントと一致しない可能性が懸念されるようになって来ている。

クリエイティブ・コモンズのほかにもオープン化を国際的に推進する団体は存在しているが、その一つ、オープンデータの分野で特に 2010 年代に影響力を持ったオープン・ナレッジ・ファウンデーションがある。中心となる組織は本拠を米国ではなく英国に置いている点が異なるが、国際的にネットワークを作つて多くの国の個人や団体・法人が連携している点ではクリエイティブ・コモンズとも似ている。オープン・ナレッジ・ファウンデーションもある時期から公正さ (fair) を重視する価値のひとつとして打ち出すようになっている。こちらは欧州が拠点であるだけに米国の巨大テック企業への警戒感がより早期に強くなつたせいかも知れない。わずかな徵候ではあるが、こうした変化は、オープン化の概念が持つ射程のようなものを検討する上でも、情報社会とオープン化の関係を整理する上でも興味深いものであるように思われる。

もっとも、既に述べたように生成 AI をめぐる状況はこのような図式でとらえてしまうと単純化が過ぎるように思われる。というのも、学習データを生み出すクリエイターとそれを利用する生成 AI 開発企業の間にはそのような構図が成り立つ場合があるものの、生成 AI の利用者まで含めて考えるならば、従来は作り手の側に参加できなかつたような多くの人が、新しいツールの助けを借りて表現に参加できる可能性もあるためだ。この部分はむしろ古典的なエンパワーメントであり、オープン性の高さ (学習データ利用の自由さを含む) がこれを後押しするという構図があるのは第 2 節で見た通りだ。また、学習データを提供する側にも大企業が存在し、AI を開発する側にも小規模な組織が存在している。

著作権制度をめぐる議論の中では、しばしば、産業として成り立っているコンテンツ関連の企業や権利者団体は自分達の利益を拡大する政策を実現するために取り組むメリットが高いが、利用者はそのような政策によってデメリットを被る度合いが小さいために、狭い範囲の権利者に集中しているメリットが、幅広い利用者に薄く存在しているメリットよりも優先されがちであるといった指摘がされる。オープン化の運動は、このような薄く広く存在している利用者のメリットに目を向ける効果を持つことが多かつたわけだが、同時にそれはオンラインで活躍するアマチュアなども含む多種多様なクリエイターの賛同・協力を得られるから成立する取り組みでもあった。

今回は部分的に似た、しかし別の構図があると言えるように思われる。クリエイティブ・コモンズは、クリエイターの意向に耳を傾けつつも、「薄く広く存在している生成 AI 利用者のメリット」にも目を向ける効果を持っているように思われる。著作権制度も根本的な改革や廃止をするわけではなく、自主的な秩序形成を通じた解決を目指す、という形は同じであり、ファイル共有などが話題になつてゐた当時にあつたような著作権廃止論を採用しないのではなく、著作権法を拡張しても既存のクリエイターの利益を守るという路線を採用しない形だ。このようなステークホルダーの構図の違いは、単に本稿の主題についてのみ

存在している特殊な事情ではなく、情報社会の変容とも呼ぶべき変化であり、他の文脈、他のイシューをめぐっても似たような形で登場することがある特徴のように思われる。新しい構図についての過度の単純化が判断を誤る可能性については既に述べた通りだが、ではどのような把握が望ましいのかについてはまた別の機会の検討の課題としたい。

5. おわりに

本稿は、生成AIの学習データをめぐるガバナンスをテーマとして、主に日本の現行の著作権制度によって実現している学習データの利用しやすさに着目して考察を行った。著作権法は、他人の著作物を無断で利用することを原則として違法としており、利用にあたっては権利者からの許諾の取得を義務付ける制度である。日本では学習データ利用についてはこの原則に対する例外が設けられており、かなり自由に利用ができるようになっている。対価の支払いなども特に義務付けられていない。これはデータをめぐる取引コストや許諾料の形で払う料金を含め、生成AI開発のコストを大きく低減する効果がある。オープンデータやオープンソースソフトウェアなどの文脈で用いられるような「オープン」という概念に照らせば、これはAIの開発に様々な人が参入・参加できる余地を作るものであり、安価なサービス提供を通じて幅広い利用者の利用も後押しする効果が期待でき、そこから更に開発の方向性・満たすべきニーズなどが見える可能性もあり、生成AIの開発と利用両面でメリットがある。

生成AIは業務利用に限っても、幅広いセクターにかなり大きなメリットをもたらすことが期待できる。ただし、そのメリットは誰もが平等に享受できるものではなく、デメリットもあり、そのデメリットの中には特定の職業や業界の市場規模が縮小するといったものも含まれると思われる。このデメリットもまた誰もが平等に被るものではない。AIの利用から得られるメリットが十分に大きいと期待されることから、基本的にはデメリットが大きい人などへの支援・救済策を導入しつつも、生成AIの開発・普及を後押しすることが望ましいと思われる。ただし、こうした考え方はAI技術の特徴や競争の構図、経済安全保障上のリスクや軍事利用などのリスク、特定の人々が大きな損害を受け、特定の人々が大きな恩恵を受けるかも知れないことについての倫理的な問題などについてどのように考えるか、またどのような前提をおくか、によっても変わってくる部分がある。

特にデメリットを被る人々への救済・支援策に関しては、データ利用の原則禁止（有料取引）、オプトアウト、補償金制度など著作権者に何らかの形で金銭的支払いが行く仕組みが議論されることも多い。このような制度は、中小のAI開発者が負担できる額を大きく超え、大手開発者を有利にする可能性があること、そのリスクを抑制しようとするとクリエイターにとって生計を立てられるような十分な支払い額にならないことなどから、望ましい解決策ではないと考えた。また、他にも制度設計上難しい点があるように思われた。例えば生成AI以外のAIに学習データを提供する人も、同じように所得の低下や雇用の減少などを経験するリスクがある場合が考えられるが、著作権制度でこうした人々を救済することにも困難がある。補償金などは受け取る資格がある者の判別なども困難が伴う。リスクリング支援など、著作権者に限らず幅広い人を支援する制度はこうした難点を考えると、望ましいものと思われる。

国際的な議論を通じて、強制力を伴わないものの、自主的にAI開発者が利益を還元する

ことを促す仕組みを作ろうという動きがある。うまく機能するかどうかは現段階では不明だが、強制力を伴わないことから AI 開発を遅延させる効果を抑制しつつ、利益還元を何らかの形で行うことからある程度の割合のクリエイターのインセンティブも大きくは削がない、といった帰結をもたらせる可能性もあるように思われる。

実際には当面のところはこのような自主的な仕組みのトライアルや、大手データ保守者からの個別契約によるデータの購入や、一部の国でのオプトアウトの制度化など、様々な取り組みが混合した状態が続き、技術や国際競争やビジネスモデル上の変化なども続くことになるだろう。その中で日本にとってより望ましい制度はどのようなものかを見通すのは容易ではないようと思われる。本稿では可能な限り主要な論点に触れつつ、平易な検討・整理を試み、この課題の難しさも併せて浮かび上がらせることに努めたつもりである。より巨視的に見ると、データの利用を容易にすることがオープン化の概念に照らせば望ましい反面、それが従来影響力の弱かった多様な個人などをエンパワーし、そのエンパワーメントが大きなメリットを生み出すことにつながる、というオープン化の典型的な構図が成り立つとは限らない面もある点が、検討を難しくしている一つの大きな理由であると思われた。これは情報社会の大きな構図の変化とも言える面があるように思われたため、特にとりあげて論じた。

筆者の様々な限界に制約を受け、粗削りなところもある議論となつたが、学際的な検討が引き続き必要な課題であり、今後の議論に貢献できるような一歩となることを願いつつ本稿を終えたい。

参考文献

1. Acemoglu, D. (2024). The simple macroeconomics of AI. *Economic Policy*, 40(121), 13-58. <https://doi.org/10.1093/epolic/eiae042>
2. Allison, D. M., & Herzog, S. (2025). Artificial Intelligence and nuclear weapons proliferation: The technological arms race for (in) visibility. *Risk Analysis*, 00:1-19. <https://doi.org/10.1111/risa.70105>
3. anonymous (2015). Definition of free culture works. <https://freedomdefined.org/Definition>
4. Artificial Analysis (2025). Artificial Analysis state of AI Q2 2025 highlights report. <https://artificialanalysis.ai/downloads/state-of-ai/2025/Q2-2025-Artificial-Analysis-State-of-AI-Highlights-Report.pdf>
5. Anthropic (2025). Threat intelligence report: August 2025. <https://www-cdn.anthropic.com/b2a76c6f6992465c09a6f2fce282f6c0cea8c200.pdf>
6. Algorithms and authoritarians: Why U.S. AI must lead: House Select Committee on the Strategic Competition Between the United States and the Chinese Communist Party, 119th Congress, 2025. <https://www.congress.gov/event/119th-congress/house-event/118428>
7. Ayers, J. W., Poliak, A., Dredze, M., Leas, E. C., Zhu, Z., Kelley, J. B., Faix, D. J., Goodman, A. M., Longhurst, C. A., Hogarth, M. & Smith, D. M. (2023). Comparing

- physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. *JAMA Internal Medicine*, 183(6), 589-596.
<https://doi.org/10.1001/jamainternmed.2023.1838>
8. Barcott, B. (2024). AI lawsuits worth watching: A curated guide, TechPolicy Press, 2024.07.02. <https://www.techpolicy.press/ai-lawsuits-worth-watching-a-curated-guide/>
 9. Berg, J., Gmyrek, P., Licata, M., Gwenyaya, T. Scaiano, F., & Sánchez, M. (2025a). Generative AI and the media and culture industry, ILO brief, International Labour Organization. <https://doi.org/10.54394/FOSL3256>
 10. Berg, J., Kamiński, K., Konopczyński, F., Ładna, A., Rosłaniec, K. & Troszyński, M. (2025b). Generative AI and jobs: A refined global index of occupational exposure, International Labour Organization working papers. <https://webapps.ilo.org/static/english/intserv/working-papers/wp140/index.html>
 11. Boiko, D. A., MacKnight, R., & Gomes, G. (2023). Emergent autonomous scientific research capabilities of large language models. arXiv preprint arXiv:2304.05332
 12. Brabham, D. C. (2008). Moving the crowd at iStockphoto: The composition of the crowd and motivations for participation in a crowdsourcing application. *First Monday*, 13(6).
<https://firstmonday.org/ojs/index.php/fm/article/download/2159/1969>
 13. Brockman, G., Sutskever, I. & OpenAI (2015). Introducing OpenAI, 2015.12.11. <https://openai.com/index/introducing-openai/>
 14. Brynjolfsson, E. Chandar, B. & Chen, R. (2025). Canaries in the coal mine?: six facts about the recent employment effects of Artificial Intelligence. <https://digitaleconomy.stanford.edu/publications/canaries-in-the-coal-mine/>
 15. Budapest Open Access Initiative (2022). Read the declaration.
<https://www.budapestopenaccessinitiative.org/read/>
 16. Byju, A. (2024). The “Godfather of AI” predicted I wouldn’t have a job. He was wrong. New Republic, 2024.10.25. <https://newrepublic.com/article/187203/ai-radiology-geoffrey-hinton-nobel-prediction>
 17. Cazzaniga, M., Jaumotte, M.F., Li, L., Melina, M.G., Panton, A.J., Pizzinelli, C., Rockall, E.J. & Tavares, M.M.M. (2024). Gen-AI: Artificial intelligence and the future of work. IMF Staff Discussion Note SDN2024/001. International Monetary Fund. <https://www.imf.org/en/Publications/Staff-Discussion-Notes/Issues/2024/01/14/Gen-AI-Artificial-Intelligence-and-the-Future-of-Work-542379>
 18. Chedraoui, K. (2025). Best AI Image generators of 2025. CNET, 2025.08.03. <https://www.cnet.com/tech/services-and-software/best-ai-image-generators/>
 19. Clark, K. (2025). Mistral set for \$14 billion (€12 billion) valuation with new funding round. Bloomberg, 2025.09.03.
<https://www.bloomberg.com/news/articles/2025-09-03/mistral-set-for-14-billion->

- valuation-with-new-funding-round
20. Clifford, S. (2010). For photographers, the image of a shrinking path. New York Times, 2010.03.30.
<https://www.nytimes.com/2010/03/30/business/media/30photogs.html>
 21. commoncrawl (undated). Size of Common Crawl monthly archives. GitHub.
<https://commoncrawl.github.io/cc-crawl-statistics/plots/crawlsize>
 22. Council of Economic Advisors (2024). An economic framework for understanding Artificial Intelligence, in Ch.7, The 2024 economic report of the president. The White House, 2024.03.21. <https://bidenwhitehouse.archives.gov/cea/written-materials/2024/03/21/the-2024-economic-report-of-the-president/>
 23. Creative Commons (undated) CC Signals: A new social contract for the age of AI.
<https://creativecommons.org/ai-and-the-commons/cc-signals/>
 24. De Cremer, D., Bianzino, N. M., & Falk, B. (2023). How generative AI could disrupt creative work. Harvard Business Review, 13, 13
 25. de Vries, M., Kapff, L., Achiaga, M. N., Wauters, P., Osimo, D., Foley, P., Szkuta, K., O'Connor, J., & Whitehouse, D. (2011). Pricing Of Public Sector Information study: Models of supply and charging for public sector information. Report for European Commission, October 2011. <https://op.europa.eu/en/publication-detail/-/publication/d0862188-b657-40b4-b3c0-f3c05d454247/language-en>
 26. Dent, S. (2024). More news organizations sue OpenAI and Microsoft over copyright infringement: The Intercept, Raw Story and AlterNet claim ChatGPT uses 'verbatim' content without attribution. Engadget, 2024.02.29.
<https://www.engadget.com/more-news-organizations-sue-openai-and-microsoft-over-copyright-infringement-061103178.html>
 27. Dunn, J. E. (2025). Swiss launch open source AI model as “ethical” alternative to big US LLMs. Info World, 2025.09.04.
<https://www.infoworld.com/article/4051779/swiss-launch-open-source-ai-model-as-ethical-alternative-to-big-us-langs.html>
 28. European Commission (2025). The General-purpose AI code of practice.
<https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai>
 29. Force11 (2023). The FAIR data principles. <https://force11.org/info/the-fair-data-principles/>
 30. Freedman, D. (2025). The day ChatGPT went cold: When OpenAI released a new version of ChatGPT, people were quick to protest its colder responses. Acknowledging the emotional attachment with chatbots, the company quickly backtracked. New York Times, 2025.08.31.
<https://www.nytimes.com/2025/08/19/business/chatgpt-gpt-5-backlash-openai.html>
 31. Free Software Foundation (2024). Free software definition.
<https://www.gnu.org/philosophy/free-sw.en.html>
 32. Frey, C. B., & Osborne, M. A. (2013). The future of employment: How susceptible

- are jobs to computerisation?. Oxford Martin School working paper. <https://oms-www.files.svcdn.com/production/downloads/academic/future-of-employment.pdf>.
33. Frey, C. B. & Osborne, M. (2024). Generative AI and the future of work: A reappraisal, *The Brown Journal of World Affairs*, 30(1). <https://bjwa.brown.edu/30-1/generative-ai-and-the-future-of-work-a-reappraisal/>
 34. Gao, L., Biderman, S., Black, S., Golding, L., Hoppe, T., Foster, C., Phang, J., He, H., Thite, A., Nabeshima, N., Presser, S., & Leahy, C. (2020). The Pile: An 800GB dataset of diverse text for language modeling. arXiv preprint arXiv:2101.00027.
 35. Global Partnership on Artificial Intelligence (2025). AI openness: A primer for policymakers, OECD Artificial Intelligence papers no. 44, 2025.08.14. <https://doi.org/10.1787/02f73362-en>
 36. Gmyrek, P., Berg, J. & Bescond, D. (2023). Generative AI and jobs: A global analysis of potential effects on job quantity and quality. ILO working paper 96, August 2023. <https://doi.org/10.54394/FHEM8239>
 37. Griffith, E. (2025). Tested: The best AI image generators for 2025. PC Magazine, 2025.01.31. <https://www.pc当地.com/picks/the-best-ai-image-generators>
 38. Hahnel, M., Graham, S. & Campbell, A. (2024). The state of open data 2024: Special report bridging policy and practice in data sharing. Digital Science, 2024. <https://doi.org/10.6084/m9.figshare.27337476.v2>
 39. Heath, A. (2025). I talked to Sam Altman about the GPT-5 launch fiasco. Verge, 2025.08.15. <https://www.theverge.com/command-line-newsletter/759897/sam-altman-chatgpt-openai-social-media-google-chrome-interview>
 40. Hinton, G. (2016). untitled comment. the 2016 Machine learning and market for intelligence conference in Toronto, Rotman School of Management, University of Toronto. <https://www.youtube.com/watch?v=2HMPRXstSvQ>
 41. Intellectual Property Office (2022). Consultation outcome: Artificial Intelligence and intellectual property: copyright and patents: Government response to consultation, 2022.06.28. <https://www.gov.uk/government/consultations/artificial-intelligence-and-ip-copyright-and-patents/outcome/artificial-intelligence-and-intellectual-property-copyright-and-patents-government-response-to-consultation>
 42. jar (2024). Topic: Adobe Stock 2024 contributor bonus, MicrostockGroup. <https://www.microstockgroup.com/fotolia.com/adobe-stock-2024-contributor-bonus/>
 43. Jeffries, S. (2013). The death of photography: are camera phones destroying an artform?, Guardian, 2013.12.13. <https://www.theguardian.com/artanddesign/2013/dec/13/death-of-photography-camera-phones>
 44. Kim, E. (2023). Internal documents show Adobe wrestling with how to spot AI images, how to pay creators for training data, and a possible opt-out. Business Insider, 2023.08.23. <https://www.businessinsider.com/adobe-firefly-ai-images-compensate-creators-model-training-data-2023-8>

45. Klyman, K. & Piliero, R. (2024). AI and the A-bomb: What the analogy captures and misses. *Bulletin of the Atomic Scientists*, 2024.09.09.
<https://thebulletin.org/2024/09/ai-and-the-a-bomb-what-the-analogy-captures-and-misses/>
46. Lee, P. (2025). Synthetic data and the future of AI. *Cornell Law Review*, 110, 1-74.
<https://publications.lawschool.cornell.edu/lawreview/2025/03/18/synthetic-data-and-the-future-of-ai/>
47. LM Arena (undated). <https://lmarena.ai/leaderboard/text-to-image>
48. Lucchi, N. (2025). Generative AI and copyright: Training, creation, regulation. Policy Department for Justice, Civil Liberties and Institutional Affairs, Directorate-General for Citizens' Rights, Justice and Institutional Affairs, PE 774.095.
[https://www.europarl.europa.eu/thinktank/en/document/IUST_STU\(2025\)774095](https://www.europarl.europa.eu/thinktank/en/document/IUST_STU(2025)774095)
49. Lutkevich, B. & Rosa Heaton, R. (2025). AI lawsuits explained: Who's getting sued?: Authors, artists and others are filing lawsuits against generative AI companies for using their data in bulk to train AI systems without permission. TechTarget, 2025.07.07, <https://www.techtarget.com/WhatIs/feature/AI-lawsuits-explained-Whos-getting-sued>.
50. Matthews, D. (2023). AI is supposedly the new nuclear weapons—but how similar are they, really?. Vox, 2023.06.29. <https://www.vox.com/future-perfect/2023/6/29/23762219/ai-artificial-intelligence-new-nuclear-weapons-future>
51. Mariella Moon (2025). Switzerland launches its own open-source AI model: Apertus was trained only on publicly available data, it said. Engadget. 2025.09.02. <https://www.engadget.com/ai/switzerland-launches-its-own-open-source-ai-model-133051578.html>
52. Maslej, N., Fattorini, L., Perrault, R. Gil, Y., Parli, V., Kariuki, N., Capstick, E., Reuel, A., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Niebles, J. C., Shoham, Y., Wald, R., Walsh, T., Hamrah, A., Santarasci, L., Lotufo, J. B., Rome, A., Shi, A. & Oak., S. (2025). The AI index 2025 annual report. AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, April 2025.
https://hai.stanford.edu/assets/files/hai_ai_index_report_2025.pdf
53. New Zealand Government (undated). The ANZ Truckometer.
<https://www.data.govt.nz/catalogue-guide/showcase/anz-truckometer>
54. Nuñez, M. (2024). Stable Diffusion creators launch Black Forest Labs, secure \$31M for FLUX.1 AI image generator. Venture Beat, 2024.08.01.
<https://venturebeat.com/ai/stable-diffusion-creators-launch-black-forest-labs-secure-31m-for-flux-1-ai-image-generator>
55. OpenAI (2025a). Introducing GPT-OSS. 2025.08.06.
<https://openai.com/index/introducing-gpt-oss/>

56. OpenAI (2025b). Disrupting malicious uses of AI. June 2025.
<https://cdn.openai.com/threat-intelligence-reports/5f73af09-a3a3-4a55-992e-069237681620/disrupting-malicious-uses-of-ai-june-2025.pdf>
57. Open Knowledge Foundation (undated). Open Definition.
<https://opendefinition.org/>
58. Open Source Initiative (2024). Open Source Definition, <https://opensource.org/osd>
59. Open Source Initiative (2025). The Open Source AI Definition.
<https://opensource.org/ai>
60. Organization for Economic Co-operation and Development (2024). Job creation and local economic development 2024: The geography of generative AI. 2024.11.28.
https://www.oecd.org/en/publications/job-creation-and-local-economic-development-2024_83325127-en.html
61. Ortiz, S. (2025). The best AI image generators are getting scary good at things they used to be terrible at these 10 best AI text-to-image generators create more realistic images than ever (and most of them are free). ZDNet, 2025.05.09.
<https://www.zdnet.com/article/best-ai-image-generator/>
62. Paul, K. & Tong, A. (2024). Inside Big Tech's underground race to buy AI training data. Reuters, 2024.04.06. <https://www.reuters.com/technology/inside-big-techs-underground-race-buy-ai-training-data-2024-04-05/>
63. Peukert, C. (2025). Copyright and the dynamics of innovation in artificial intelligence. Proceedings of the 58th Hawaii International Conference on System Sciences. <https://doi.org/10.24251/HICSS.2025.538>
64. PMP Consulting (2024). Study on the economic impact of Generative AI in the Music and Audiovisual industries Complete study: Current situation and 5-year perspective. International Confederation of Societies of Authors and Composers, November 2024. <https://www.cisac.org/Newsroom/news-releases/global-economic-study-shows-human-creators-future-risk-generative-ai>
65. Ramesh, A., Dhariwal, P., Nichol, A., Chu, C. & Chen, M. (2022). Hierarchical text-conditional image generation with CLIP latents.
<https://doi.org/10.48550/arXiv.2204.06125>
66. Ramesh, A., Pavlov, M., Goh, G., Gray, S., Voss, C., Radford, A., Chen, M. and Sutskever, I. (2021). Zero-shot text-to-image generation. Proceedings of the 38th International Conference on Machine Learning, PMLR 139, 8821-8831.
<http://proceedings.mlr.press/v139/ramesh21a/ramesh21a.pdf>
67. Roehrlich, E. (2025). Why is the nuclear-AI analogy so popular?. *Nature Reviews Physics*, 1-2. <https://www.doi.org/10.1038/s42254-025-00873-x>
68. Rombach, R., Blattmann, A., Lorenz, D., Esser, P. & Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 10684-10695.
69. Rombach, R. & Esser, P. (2022). Stable Diffusion v1-1 model card,

- <https://huggingface.co/CompVis/stable-diffusion-v1-1>
70. Saharia, C., Chan, W., Saxena, S., Li, L., Whang, J., Denton, E. Ghasemipour, S. K. S., Ayan, B. K., Mahdavi, S.S., Gontijo-Lopes, R., Slimans, T., Ho, J., Fleet, D. J. & Norouzi, M. (2022). Photorealistic text-to-image diffusion models with deep language understanding. The 36th Conference on Neural Information Processing Systems (NeurIPS 2022).
https://proceedings.neurips.cc/paper_files/paper/2022/hash/ec795aeadae0b7d230fa35cbaf04c041-Abstract-Conference.html
71. Samberg, R., Vollmer, T., & Padilla, T. (2023). Legal literacies for text data mining, Crossborder (“LLTDM-X”) White Paper.
<https://escholarship.org/uc/item/5k91r1s1>
72. Schuhmann, C., Beaumont, R., Vencu, R., Gordon, C., Wightman, R., Cherti, M., Coombes, T., Katta, A., Mullis, C., Wortsman, M., Schramowski, P., Kundurthy, P., Crowson, K., Schmidt, L., Kaczmarczyk, R. & Jitsev, J. (2022). LAION-5B: An open large-scale dataset for training next generation image-text models. The 36th Conference on Neural Information Processing Systems (NeurIPS 2022).
<https://arxiv.org/abs/2210.08402>
73. Service, R. F. (2023). Could chatbots help devise the next pandemic virus?. Science, 380(6651), 1211. <https://doi.org/10.1126/science.adj2463>
74. Shcherbakov, V., Dalaud, I., & Peukert, C. (2025). AI needs better data than the law allows. SSRN 5296505. <http://doi.org/10.2139/ssrn.5296505>
75. Shumailov, I., Shumaylov, Z., Zhao, Y., Papernot, N., Anderson, R., & Gal, Y. (2024). AI models collapse when trained on recursively generated data. *Nature*, 631(8022), 755-759. <https://doi.org/10.1038/s41586-024-07566-y>
76. Sigalos M. (2025). OpenAI raises \$8.3 billion as paid ChatGPT business users reach 5 million. CNBC, 2025.08.01. <https://www.cnbc.com/2025/08/01/openai-raise-chatgpt-users.html>
77. Sorkin, A. R., Warner, B., Kessler, S., de la Merced, M. J., & Kaye, D. (2025). OpenAI raises another funding deal, from Dragoneer, Blackstone and more. New York Times, 2025.08.04.
<https://www.nytimes.com/2025/08/01/business/dealbook/openai-ai-mega-funding-deal.html>
78. Sriram, A. & Hu, K. (2024). Cash-strapped Stability AI raises \$80 mln with new CEO and board. Reuters, 2024.06.25.
<https://www.reuters.com/technology/artificial-intelligence/cash-strapped-stability-ai-raises-80-mln-with-new-ceo-board-2024-06-25/>
79. Stern, J. (2023). AI Is Like ... Nuclear Weapons?: The new technology is beyond comparison. The Atlantic, 2023.03.26.
<https://www.theatlantic.com/technology/archive/2023/03/ai-gpt4-technology-analogy/673509/>

80. Taub, E. A. (2007). When are photos like penny stocks? When they sell. New York Times, 2007.06.05.
<https://www.nytimes.com/2007/06/05/technology/circuits/05syndicate.html>
81. The White House (2025). America's AI action plan: Winning the race. 2025.07.
<https://www.whitehouse.gov/wp-content/uploads/2025/07/Americas-AI-Action-Plan.pdf>
82. Tong, A., Wang, E. & Coulter, M. (2024). Exclusive: Reddit in AI content licensing deal with Google. Reuters, 2024.02.22. <https://www.reuters.com/technology/reddit-ai-content-licensing-deal-with-google-sources-say-2024-02-22/>
83. Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M. A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., Rodriguez, A., Joulin, A., Grave, E., & Lample, G. (2023). Llama: Open and efficient foundation language models. arXiv preprint arXiv:2302.13971.
84. Trimble, M. (2025). Cross-border limitations and exceptions to copyright: "Powered by AI." Joint PIJIP/TLS Research Paper Series, 151.
<https://digitalcommons.wcl.american.edu/research/151>
85. Vilendreer, E. K. (2025). Update on generative AI litigation: A future hanging in the balance. American Bar Association, 2025.05. 09.
https://www.americanbar.org/groups/science_technology/resources/scitech-lawyer/2025-spring/generative-ai-litigation-future-update/
86. Weisenberger, T. M., Milton, D. C. & Enright, H. A. (undated). Case tracker: Artificial Intelligence, copyrights and class actions. BakerHostetler.
<https://www.bakerlaw.com/services/artificial-intelligence-ai/case-tracker-artificial-intelligence-copyrights-and-class-actions>
87. Widder, D. G., Whittaker, M. & West, S. M. (2024). Why 'open' AI systems are actually closed, and why this matters. *Nature* 635, 827–833.
<https://doi.org/10.1038/s41586-024-08141-1>
88. Wright, W. (2025). You can track the top AI image generators via this new leaderboard - and vote for your favorite too: Feeling overwhelmed by the number of text-to-image models that are out there to choose from? LMArena.ai shows you which ones stand out from the crowd. ZD Net, 2025.06.05.
<https://www.zdnet.com/article/you-can-track-the-top-ai-image-generators-via-this-new-leaderboard-and-vote-for-your-favorite-too/>
89. ABEJA Tech Blog (2024). 「Common Crawl から作る大規模日本語コーパスとその前処理 (Mixtral 8x7B を語彙拡張継続事前学習 Part2)」 2024.05.07. <https://tech-blog.abeja.asia/entry/abeja-nedo-project-part2-202405>
90. IT Media (2025a). 「GPT-5 の “性格” 変更へ 『温かみがありつつ、GPT-4o ほど煩わしくないものに』」 . 2025.08.13.
<https://www.itmedia.co.jp/apiplus/articles/2508/13/news099.html>
91. IT Media (2025b). 「学習データの権利クリア、画像生成AI「AC 写真 AI ラボ」登場 商用利用OK AC ワークスから」 2025.04.02.

- <https://www.itmedia.co.jp/aiplus/articles/2504/02/news160.html>
92. NBDC 研究員 (大波純一、八塚茂、信定知江、箕輪真理、三橋信孝、畠中秀樹) (2019). 『FAIR 原則 (「THE FAIR DATA PRINCIPLES」和訳)』
<https://doi.org/10.18908/a.2019112601>
93. Newman, D. (2025). 「AI はバブルかインフラか? 73 兆円評価の OpenAI が狙う数百兆円投資」 Forbes Japan, 2025.08.25.
<https://forbesjapan.com/articles/detail/81485>
94. 『NOMORE 無断生成 AI』声優有志の会 (2024). Statement. <https://nomore-mudan.com/#statement>
95. 天瀬光二、荻野登、戸田卓宏 (2025). 「AI の職場導入による働き方への影響等に関する調査 (労働者 Web アンケート) 結果」 2025.05.23.
<https://www.jil.go.jp/institute/research/2025/256.html>
96. 五十嵐輝、田中絵麻、小山友介 (2023). 『クリエイターエコノミーに関する動向調査(前編)』一般財団法人 マルチメディア振興センター.
https://www.fmmc.or.jp/Portals/0/ICT/forum/2022_creator.pdf
97. 五十嵐輝、田中絵麻、小山友介 (2024). 『クリエイターエコノミーに関する動向調査(後編)』一般財団法人 マルチメディア振興センター.
https://www.fmmc.or.jp/Portals/0/ICT/forum/2023_creator-economy.pdf
98. 池内有為 (2025). 「FAIR 原則の認知と実践に関する国際比較 : The State of Open Data 2024 調査データの二次分析 (連載 : 続・オープンサイエンスのいま)」情報の科学と技術 75(5) 234-237. https://www.jstage.jst.go.jp/article/jkg/75/5/75_234/_pdf-char/ja
99. 浦中美穂 (2024). 「サカナ AI に 3 メガバンクなど出資 一連の資金調達約 300 億円」 Reuters, 2024.09.17. <https://jp.reuters.com/markets/global-markets/AM35GW72DRJCBUALA37Z562VM-2024-09-17/>
100. 栄藤稔. (2023). 「生成 AI の発展と新たなコンテンツ経済圏」 情報通信政策研究, 7(1), 25-51. https://doi.org/10.24798/jicp.7.1_25.
101. 王威駒 (2025). 「論説: AI と法研究(2) 生成 AI とプラットフォーム事前規制」 比較法学 58 (3), 1-23. <https://waseda.repo.nii.ac.jp/records/2005347>
102. 大野敬太郎 (2016). 「知的財産戦略の提言」 <https://keitaro-ohno.com/3151/>
103. 奥邨弘司 (2025). 「生成 AI と著作権:「考え方について」に関する 3 つの論点」 知的財産法政策学研究, 70, 103-143. <http://hdl.handle.net/2115/95534>
104. 岡部敬史、市川史樹 (2025). 「勝又清和七段 将棋 AI の登場で変わったもの、復活したもの」 日経 BOOK プラス, 2025.4.15.
<https://bookplus.nikkei.com/atcl/column/031000478/031000003/>
105. 小沢高広 (2024). 『漫画制作における生成 AI 活用の現状 : 2024 春』 経済産業研究所 BBL, 2024.3.8. <https://www.rieti.go.jp/jp/events/bbl/24030801.html>
106. 小山紘一 (2023). 「著作権法の令和 5 年改正とデジタルアーカイブの課題」 デジタルアーカイブ学会誌 7(4) p. 169-172. https://doi.org/10.24506/jsda.7.4_169
107. 岳梁 (2024). 「拡大するサービス貿易と日本のデジタル赤字」 DBJ Research No.423, 日本政策投資銀行, 2024.09.09.
<https://www.dbj.jp/upload/investigate/docs/f038c03994dbf43db3edb8bc7ccaba4e.pdf>
108. 梶川智博 (2024). 「放射線治療における人工知能 (AI) の役割」 京都府立医科大学雑誌 133(2) 117-126. <https://doi.org/10.32206/jkpum.133.02.117>
109. 公正取引委員会 (2025). 『生成 AI に関する実態調査報告書 ver.1.0』

- <https://www.jftc.go.jp/houdou/pressrelease/2025/jun/250606generativeai.html>
110. 厚生労働省 (2025). 『令和6年雇用動向調査結果の概況』
<https://www.mhlw.go.jp/toukei/itiran/roudou/koyou/doukou/25·2/dl/gaikyou.pdf>
111. 古作登 (2023). 「将棋AIの進化が公式戦の戦型選択に与えた影響-プロが模倣する戦法と、ソフト評価の低い振り飛車が選ばれる理由-」 大阪商業大学アミューズメント産業研究所紀要, 25, 33-50. <https://ouc.repo.nii.ac.jp/records/2000474>
112. サイバーエージェント (2023). 「サイバーエージェント、最大68億パラメータの日本語LLM(大規模言語モデル)を一般公開 一オープンなデータで学習した商用利用可能なモデルを提供」, 2023.05.17.
<https://www.cyberagent.co.jp/news/detail/id=28817>
113. 柿原美紀 (2016). 『「柔軟性のある規定」の導入の必要性について』,
https://home.jeita.or.jp/press_file/20200121155521_kNodU8Zt9w.pdf
114. 自由民主党政務調査会 (2016). 『地方創生とイノベーション創出のための知的財産戦略 提言: 第4次産業革命とグローバル化の中で』 2016.04.21.
https://storage2.jimin.jp/pdf/news/policy/132110_1.pdf
115. ジェームズ・スロウイッキー (2006). 『「みんなの意見」は案外正しい』 (小高尚子訳、角川書店)
116. スコット・ペイジ (2009). 『「多様な意見」はなぜ正しいのか 衆愚が集合知に変わること』 (水谷淳訳、日経BP)
117. 数藤雅彦・橋本阿友子 (2019). 『保護期間満了(パブリックドメイン)の判断基準』『権利処理と法の実務(デジタルアーカイブ・ベースิกス 1)』(福井建策(監修)、数藤雅彦(責任編集)) 勉誠出版、p.17-47.
118. 総務省 (2023). 『令和4年就業構造基本調査: 結果の概要』
<https://www.stat.go.jp/data/shugyou/2022/pdf/kgaiyou.pdf>
119. 総務省 (2024). 『情報通信白書 令和6年度版』
<https://www.soumu.go.jp/johotsusintokei/whitepaper/r06.html>
120. 総務省 (2025). 『情報通信白書 令和7年度版』
<https://www.soumu.go.jp/johotsusintokei/whitepaper/r07.html>
121. ソフトウェア情報センター (2024). 『ソフトウェア等の権利保護に関する調査研究報告書: 2023(令和5)年度』
<https://www.softic.or.jp/application/files/3617/2361/7601/2023R5-1.pdf>
122. 田中絵麻 (2023). 『AI時代におけるクリエイター・エコノミーの展望』 情報通信学会誌 41(2) 65-71. https://www.jstage.jst.go.jp/article/jsicr/41/2/41_65/_article/-char/ja
123. 田中辰雄 (2018). 『著作権集中管理団体の功罪をめぐる論争について: JASRACの「音楽教室からの料金徴収問題」を題材に』 知的財産法政策学研究, 51, 65-112.
<http://hdl.handle.net/2115/72119>.
124. 知財戦略本部 構想委員会 コンテンツ小委員会 デジタル時代における著作権制度・関連政策の在り方検討タスクフォース(2021). 『中間とりまとめ』
https://www.kantei.go.jp/jp/singi/titeki2/tyousakai/kousou/digital_kentou_tf/pdf/tyukan_torimatome.pdf
125. 津田通隆, 松尾武将, 栗原涼介, 岡口正也, 小林うらら, 中井正志, 羽原健雄 (2025). 『デジタル経済レポート: データに飲み込まれる世界、聖域なきデジタル市場の生存戦略』, 経済産業省大臣官房若手新政策プロジェクトPIVOT レポート.
https://www.meti.go.jp/policy/it_policy/statistics/digital_economy_report.html
126. デジタル庁 (2025). 『オープンデータ100』
https://www.digital.go.jp/resources/data_case_study
127. 内閣府 政策統括官(経済財政分析担当) (2024). 『AIで変わる労働市場』, 第1章, 『世界経済の潮流』, https://www5.cao.go.jp/j-j/sekai_chouryuu/sh24-01/index-pdf.html

128. 中田典生. (2024). 「医療分野における AI 革命: 画像診断支援からマルチモーダル大規模言語モデルへの展望」. 総合健診, 51(2) 224-228.
<https://doi.org/10.7143/jhep.51.224>
129. 西角直樹, 綿谷謙吾 (2024). 「デジタル赤字拡大は悪いことなのか?: 目指すべきは「日本の強み」と「デジタル」の融合」 三菱総合研究所, 2024.04.25.
<https://www.mri.co.jp/knowledge/insight/20240425.html>
130. ニック・ボストロム (2017). 『スーパーインテリジェンス超絶 AI と人類の命運』 (倉骨彰訳、日経 BP)
131. 新田堯之 (2024). 「生成 AI が描く日本の職業の明暗とその対応策: AI と職業情報を活用した独自のビッグデータ分析」 大和総研調査季報= DIR research quarterly review, 54, 38-71.
https://www.dir.co.jp/report/research/economics/japan/20240425_030145.html
132. 日本アニメフィルム文化連盟 (2023). 『アニメ業界を対象とした生成 AI に関する意識調査の結果』 PR Times, 2023.09.12.
<https://prtimes.jp/main/html/rd/p/000000002.000121993.html>
133. 日本音楽著作権協会 (2023). 『生成 AI と著作権の問題に関する基本的な考え方』, プレスリリース, 2023.07.24,
https://www.jasrac.or.jp/information/release/23/07_3.html
134. 日本音楽著作権協会 (2024). 『AI に関する音楽団体協議会』の設置について: for Creators, for Artists』 2024.01.25.
https://www.jasrac.or.jp/information/release/24/01_3.html
135. 日本脚本家連盟, 日本シナリオ作家協会 (2023). 『脚本家の権利と生成 AI に関する共同声明: 2023 年 11 月 3 日文化の日に』 2023.11.03. <https://mitte-x-img.istsw.jp/scenario/file/kyodoseimei20231103.pdf>
136. 日本雑誌協会, 日本写真著作権協会, 日本書籍出版協会, 日本新聞協会 (2023). 『生成 AI に関する共同声明』 2023.08.17.
https://www.pressnet.or.jp/statement/copyright/230817_15114.html
137. 日本写真家協会 (2023). 『生成 AI 画像についてその考え方の提言』
<https://www.jps.gr.jp/about-generated-ai-images/>
138. 日本美術著作権連合 (2023). 『生成 AI に関する声明』 2023.08.17.
https://www.jart.tokyo/jart_wp/wp-content/uploads/2023/08/5bbd651f0394f60cbdbf80279ca1b5ac.pdf
139. 日本貿易振興機構 市場開拓・展示事業部 海外市場開拓課 ニューヨーク事務所 (2022). 『米国におけるデジタルヘルス市場動向調査』
https://www.jetro.go.jp/ext_images/_Reports/02/2022/cba9066cb65f4ce5/202203.pdf
140. 幡谷龍一郎. (2024). 「生成データが今後のデータセットに与える影響について」 第 38 回 人工知能学会全国大会論文集, 3K5OS2b01. 人工知能学会.
https://doi.org/10.11517/pjsai.JSAI2024.0_3K5OS2b01
141. 深澤匠(2024). 「絵書きと生成 AI」 紙パ技協誌 78(10) 896
https://www.jstage.jst.go.jp/article/jtappij/78/10/78_78.896/_article/-char/ja/
142. 福岡真之介 (2022). 「AI と知的財産権 (連載 実践 知財法務 第 12 回)」 ジュリスト, 2022 年 10 月号 (No.1576), pp.86-92.
143. 文化審議会著作権分科会法制度小委員会 (2024). 『AI と著作権に関する考え方について』
https://www.bunka.go.jp/seisaku/bunkashingikai/chosakuken/pdf/94037901_01.pdf
144. 文化庁 (2018a). 『著作権法の一部を改正する法律 概要説明資料』
https://www.bunka.go.jp/seisaku/chosakuken/hokaisei/h30_hokaisei/pdf/r1406693_02.pdf
145. 文化庁 (2018b). 『著作権法の一部を改正する法律 (平成 30 年改正) について (解

- 説)』 ,
https://www.bunka.go.jp/seisaku/chosakuken/hokaisei/h30_hokaisei/pdf/r1406693_11.pdf
146. 文化庁著作権課 (2019). 『デジタル化・ネットワーク化の進展に対応した柔軟な権利制限規定に関する基本的な考え方 (著作権法第30条の4, 第47条の4及び第47条の5関係)』
https://www.bunka.go.jp/seisaku/chosakuken/hokaisei/h30_hokaisei/pdf/r1406693_17.pdf
147. 文化庁著作権課 (2025). 「2026年度から始まる未管理著作物裁定制度について」 カレントアウェアネス-E, E2800, No.504, 2025.07.03. <https://current.ndl.go.jp/e2800>
148. 前田健 (2024). 「V生成AIの利用が著作権侵害となる場合 (◆特集 AIから法の世界)」 法学教室 2024年4月号(No.523), pp.27-31.
149. 町田幹 (2023). 「画像診断AIへの期待」 日本医科大学医学会雑誌, 19(3) 286-290.
<https://doi.org/10.1272/manms.19.286>
150. 松井優二郎、文谷和磨 (2022). 「わが国ベンチャービジネスの現状と課題」 日銀レビュー, 2022-J-11. https://www.boj.or.jp/research/wps_rev/rev_2022/data/rev22j11.pdf
151. マックス・テグマーク (2019). 『Life 3.0: 人工知能時代に人間であるということ』 (水谷淳訳、紀伊國屋書店)
152. 三菱UFJリサーチ&コンサルティング (2024). 『国内クリエイターエコノミーに関する調査結果(2024年)』 https://www.murc.jp/library/report/cr_241216/
153. 横山健一 & 桶川隆嗣 (2021). 「対談: AIの時代における放射線科医の役割と今後の展望」 杏林医学会雑誌, 52(4), 251-255. <https://doi.org/10.11434/kyorinmed.52.251>
154. 労働政策研究・研修機構 (2025a). 「生成AIがメディア文化産業に及ぼす影響: ILO研究ブリーフ」 https://www.jil.go.jp/foreign/jihou/2025/06/ilo_01.html
155. 労働政策研究・研修機構 (2025b). 「世界の雇用の4分の1が生成AIに代替される可能性: ILO研究ブリーフ」 https://www.jil.go.jp/foreign/jihou/2025/07/ilo_02.html
156. 労働政策研究・研修機構 (2025c). 「生成AIが労働市場に与える影響を分析、地域間格差拡大の可能性: OECD『雇用創出と地域経済発展2024』」
https://www.jil.go.jp/foreign/jihou/2025/03/oecd_01.html
157. 渡邊智暉 (2010). 「ウィキペディアから「出版」を考える」 岡本真・仲俣暁生編著『ブックビジネス2.0: ウェブ時代の新しい本の生態系』
158. 渡邊智暉 (2018). 「オープンデータの価値の未知性と活用事例 (特集データ利活用と地方創生)」 J-LIS 地方自治情報誌 5(6) 39-46.